

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:32:28 ; Search time 14.28 Seconds

(without alignments)
312.814 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Perfect score: 362
Sequence: 1 MKRSVAVMLVGLSLGVPQF.....FSCPCPDGFTDPNCSVVEY 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 15437

Minimum DB seq length: 0
Maximum DB seq length: 65

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database : PIR_67:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	69.5	19.2	53	S17294	epidermal growth f
2	66	18.2	55	F44336	neurotoxin Tx3-6 -
3	62	17.1	20	A48394	major fat-globule
4	59	16.3	57	PC4415	ErbB kinase activa
5	54.5	15.1	56	P82568	hypothetical prote
6	53.5	14.8	43	A58998	epidermal growth f
7	53	14.6	63	A4484	metallothionein I
8	52	14.4	19	A31252	metallothionein I
9	51.5	14.2	61	SMRT1	metallothionein I
10	50.5	14.0	60	S30567	metallothionein I
11	49.5	13.7	64	S31825	metallothionein A
12	49	13.5	65	B25103	ferredoxin 214Fe-4
13	48.5	13.4	62	A53640	metallothionein 4
14	48	13.3	64	PERMN	ferredoxin 214Fe-4
15	47.5	13.0	64	A25775	metallothionein A
16	47	12.8	62	B53640	metallothionein IV
17	46.5	12.8	61	SMH01A	metallothionein IA
18	46.5	12.8	61	A3889	metallothionein I
19	45.5	12.6	60	T09258	metallothionein-1I
20	45.5	12.6	61	SMH02	metallothionein IE
21	45.5	12.6	61	SMH02	metallothionein II
22	45.5	12.6	61	SMH02	metallothionein I
23	45.5	12.6	61	S00808	metallothionein Ia
24	45.5	12.6	61	I46414	metallothionein-Ia
25	45.5	12.6	61	I46559	metallothionein -
26	45	12.4	26	SMNC	metallothionein -
27	45	12.4	44	I48942	cellular disintegr
28	45	12.4	61	S47651	metallothionein IH
29	45	12.4	61	S14049	metallothionein pr

30	44.5	12.3	60	JC2420	metallothionein -
31	44	12.2	38	A45495	beta-defensin-1 -
32	44	12.2	44	I48945	cellular disintegr
33	44	12.2	52	S65712	metallothionein I
34	44	12.2	58	H64377	ferredoxin 214Fe-4
35	44	12.2	60	SMH01A	metallothionein IA
36	44	12.2	60	S38335	metallothionein -
37	44	12.2	62	I51538	metallothionein -
38	43.5	12.0	42	D47753	beta-defensin-13 -
39	43.5	12.0	61	SMMSI	metallothionein I
40	43.5	12.0	61	I54574	metallothionein-1
41	43	11.9	39	A36453	decorin - leech (
42	43	11.9	46	A44794	antimicrobial pept
43	43	11.9	60	S36382	ig heavy chain V r
44	43	11.9	60	JC2419	metallothionein -
45	43	11.9	61	SMH01B	metallothionein IB

ALIGNMENTS

RESULT 1
S17294
epidermal growth factor - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S17294
R:Pascaill, J.C.; Jones, D.S.C.; Doel, S.M.; Clements, J.M.; Hunter, M.; Fallon, T.; E
J. Mol. Endocrinol. 6, 63-70, 1991
A>Title: Cloning and characterization of a gene encoding pig epidermal growth factor.
A:Reference number: S17294; M0ID:91197366
A:Accession: S17294
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-53 <PAS>
A:Cross-references: EMBL:X59516; NID:q1940; PUD:CAA42102.1; PID:q938287
C:Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor WYTD-con
F:6-42/Domain: EGF homology <EGF>

Query Match 19.2% Score 69.5; DB 2; Length 53;
Best Local Similarity 35.1%; Pred. No. 0.7;
Matches 13; Conservative 7; Mismatches 14; Indels 3; Gaps 1;

OY 26 CDPNP---CENGIGICLPGLAVGFSFCECPDGFTDPC 59
DB 6 CPPSHDGYCLRGVCWYIEAVDSYACNCVGVGEC 42

RESULT 2
F44336
neurotoxin Tx3-6 - spider (Phoneutria nigriventer)
C:Species: Phoneutria nigriventer
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 07-Feb-1997
C:Accession: F44336
R:Cordeiro M do, N.; de Figueiredo, S.G.; Valentim A do, C.; Diniz, C.R.; von Eickste
Toxicol. 31, 35-42, 1993
A>Title: Purification and amino acid sequences of six Tx3 type neurotoxins from the v
A:Reference number: A44336; M0ID:93190315
A:Accession: F44336
A:Status: preliminary
A:Molecule type: Protein
A:Residues: 1-95 <CDR>
A:Experimental source: venom
A:Note: sequence extracted from NCBI backbone (NCBIP:126777)
C:Keywords: neurotoxin

Query Match 18.2% Score 66; DB 2; Length 55;
Best Local Similarity 36.4%; Pred. No. 1.7;
Matches 12; Conservative 6; Mismatches 11; Indels 4; Gaps 1;

OY 22 KGDIDPN---PCENGICLPGLAVGFSFCEC 50

Db 5 RGEICTDDCECCGDBNCCYCPPOSSIGIFKSC 37

RESULT 3

A8394

major fat-globule membrane protein/MGF-E8 homolog - guinea pig (fragment)

C:Species: Cavia porcellus (guinea pig) 18-Nov-1994 #text_change 31-Oct-1997

C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 31-Oct-1997

C:Accession: A48394

R:Matner, I.H.; Banghart, L.R.; Lane, W.S.

Biochem. Mol. Biol. Int. 29, 545-554, 1993

A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig

II-like sequences.

A:Reference number: A48394; MUID:93250576

A:Accession: A48394

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <MAP>

A:Experimental source: milk

A:Note: sequence extracted from NCBI backbone (NCBI:131440)

C:Superfamily: milk fat globule protein; discolidin I amino-terminal homology; EGF homolo

Query Match 17.1%; Score 62; DB 2; Length 20;

Best Local Similarity 61.1%; Pred. No. 1.9;

Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 23 GDIDPNCPCNGICLPG 40

Db 3 GDFCDSSICLNGCTLEFG 20

RESULT 4

PC4415

ErDB Kinase activator beta, brain and thymus - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Dec-1997 #sequence_revision 10-Dec-1997 #text_change 11-Jan-2000

C:Accession: PC4415

R:Hisashimaya, S.; Horikawa, M.; Yamada, K.; Ichino, N.; Nakano, N.; Nakagawa, T.; Miyag

J. Biochem. 122, 675-680, 1997

A:Title: A novel brain-derived member of the epidermal growth factor family that interac

A:Reference number: JCS700; MUID:98006324

A:Accession: PC4415

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-57 <HIS>

A:Cross-references: DDBJ:DB9997; NID:92605633; PIDN:BA23346.1; PID:92605634

A:Experimental source: PC-12 cell

A:Comment: This protein is a member of the epidermal growth factor family. It is functio

ating the differentiation of MDA-MB-453 cells.

C:Keywords: unassigned EGF-related proteins; EGF homology

F:1-25/Domain: EGF homology (fragment) <EGF>

Query Match 16.3%; Score 59; DB 2; Length 57;

Best Local Similarity 32.3%; Pred. No. 9.5;

Matches 10; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

QY 35 GICLPLAVGSFSCPCPDGFTDPCNSVVEV 65

Db 1 GVCYIEGINDLSCKCPVGYTGDRCOQFAMV 31

RESULT 5

B82568

hypothetical protein XF2350 (imported) - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82568

R:Anonymous; The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406: 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: B82568

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-56 <STM>

A:Cross-references: GB:AE004045; GB:AE003849; NID:99107517; PIDN:AAE8149.1; GSPDB:GN

A:Experimental source: Strain 9a5c

R:Stimpson, A.J.G.; Reinech, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre

as-Neto, E.; Docena, C.; El-Dorri, H.; Fachinani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitaajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miraca, E.C.; Miyaki, C.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava

A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Teuhako, M.H.; Valinda, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.

A:Reference number: A59328

A:Contents: annotation

A:Gene: XF2350

Query Match 15.1%; Score 54.5; DB 2; Length 56;

Best Local Similarity 40.7%; Pred. No. 28;

Matches 11; Conservative 3; Mismatches 8; Indels 5; Gaps 1;

QY 30 PCENGI-----CLPLAVGSFSCPCP 51

Db 13 PCGGGLGVEMCLPCDVNSNCATCP 39

RESULT 6

A59398

epidermal growth factor - great pond snail (fragment)

C:Species: Lymnaea stagnalis (great pond snail)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-Jan-2000

C:Accession: A58998

R:Nagle, G.

submitted to the Protein Sequence Database, July 1999

A:Description: A molluscan epidermal growth factor is expressed both during CNS devel

A:Reference number: A58998

A:Accession: A58998

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-43 <NAG>

A:Experimental source: albumen gland

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: blocked amino end; pyroglutamic acid

F:8-39/Domain: EGF homology <EGF>

F:1/Modified site: pyrrolidone carboxylic acid (Glu) #status experimental

Query Match 14.8%; Score 53.5; DB 2; Length 43;

Best Local Similarity 36.6%; Pred. No. 28;

Matches 15; Conservative 4; Mismatches 17; Indels 5; Gaps 3;

QY 21 GKG-DICDPNCPENGICLPLAVGSFSC-CECPDGFDPNC 59

Db 2 GDGDDPCQIVRCISYGANC--IAYGPAICECPFGVSGIRC 39

RESULT 7

A34484

metallothionein I - yeast (Candida glabrata)

C:Species: Candida glabrata

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Aug-1999

C:Accession: A34484

R:Mehta, R.K.; Garey, J.R.; Butt, T.R.; Gray, W.R.; Winge, D.R.

J. Biol. Chem. 264, 19747-19753, 1989

C:Accession: A33825
 R:Harlow, P.; Watkins, E.; Thornton, R.D.; Nemer, M.
 Mol. Cell. Biol. 9, 5445-5455, 1989
 A:Title: Structure of an ectodermally expressed sea urchin metallothionein gene and char
 A:Reference number: A33825; MUID:90066495
 A:Accession: A33825
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-64 <NAR>
 A:Cross-references: GB:M30506; NID:g161545; PIDN:AAA30067.1; PID:g161546
 C:Superfamily: metallothionein

Query Match 13.7%; Score 49.5; DB 2; Length 64;
 Best Local Similarity 28.6%; Pred. No. 1e+02;
 Matches 14; Conservative 8; Mismatches 16; Indels 11; Gaps 4;

OY 20 FGKDDCPNCPENGICLPGLAVGSFSCPCPDG-----FTDPNCS 60
 DB 18 FGQ-DCKRTGECCKKGTCTC-GICITNA-ACKCANCCKGSGSCCTESNCA 63

RESULT 12
 B25103
 ferredoxin 2[4Fe-4S]-like protein - Azotobacter chroococcum
 C:Species: Azotobacter chroococcum
 C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 11-Jun-1999
 A:Accession: B25103; S14693; S10430
 R:Robson, R.; Woodley, P.; Jones, R.
 EMBO J. 5, 1159-1163, 1986
 A:Title: Second gene (nifH*) coding for a nitrogenase iron protein in Azotobacter chrooc
 A:Reference number: A91038
 A:Accession: B25103
 A:Molecule type: DNA
 A:Residues: 1-65 <ROB>
 R:Fallik, E.; Robson, R.L.
 Nucleic Acids Res. 18, 4616, 1990
 A:Title: Completed sequence of the region encoding the structural genes for the vanadium
 A:Reference number: S14692; MUID:90356423
 A:Accession: S14693
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-65 <FAL>
 A:Cross-references: EMBL:X51756; NID:938755; PIDN:CAA36056.1; PID:938757
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1990
 C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology
 C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein
 F:3-62/Domain: ferredoxin 2[4Fe-4S] homology <FER>
 F:10/13/16/34/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
 F:20/39/42/50/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 13.5%; Score 49; DB 2; Length 65;
 Best Local Similarity 39.1%; Pred. No. 1.2e+02;
 Matches 18; Conservative 3; Mismatches 19; Indels 6; Gaps 5;

OY 23 GDIDCPNCPENGICLPGLAVGSFSC-ECPPDFTDPNCSVVEV 65
 DB 14 GD-CKP-VCPYSGIVLGSGIVYDADSCNRCAD-LGEPKCLGVCY 56

RESULT 13
 A53640
 metallothionein 4 - human
 N:Alternate names: metallothionein IV
 C:Species: Homo sapiens (man)
 C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 03-Dec-1999
 A:Accession: A53640
 R:Quaife, C.J.; Findley, S.D.; Erickson, J.C.; Froelick, G.J.; Kelly, E.J.; Zambrowicz,
 Biochemistry 33, 7250-7259, 1994
 A:Title: Induction of a new metallothionein isoform (MT-IV) occurs during differentiat
 A:Reference number: A53640; MUID:94271779
 A:Accession: A53640

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-62 <QUA>
 A:Cross-references: GB:U07807; NID:9466264; PIDN:AAA20232.1; PID:9516535
 C:Genetics:
 A:Introns: 11/1; 33/1
 C:Superfamily: metallothionein

Query Match 13.4%; Score 48.5; DB 2; Length 62;
 Best Local Similarity 35.1%; Pred. No. 1.3e+02;
 Matches 13; Conservative 5; Mismatches 8; Indels 11; Gaps 3;

OY 27 DPNP-CENGICLPGLAVGSFSCPCPDGFTDPNCS 61
 DB 2 DPRECVMSGICMGCMG-----DNCKC---TTCNCKT 29

RESULT 14
 FERMN
 ferredoxin 2[4Fe-4S] - Rhizobium meliloti
 C:Species: Rhizobium meliloti
 C:Date: 21-May-1990 #sequence_revision 27-Jan-1995 #text_change 11-Jun-1999
 A:Accession: C32361; JE0034; S04411
 R:Muligan, M.E.; Bulkema, W.J.; Haselkorn, R.
 J. Bacteriol. 170, 4406-4410, 1988
 A:Title: Bacterial-type ferredoxin genes in the nitrogen fixation regions of the cyan
 A:Reference number: A91884; MUID:88314954
 A:Accession: C32361
 A:Molecule type: DNA
 A:Residues: 1-64 <KLU>
 A:Cross-references: GB:M21841; NID:9152194; PIDN:AAA26268.1; PID:9152195
 R:Klipp, W.; Reilander, H.; Schueller, A.; Krey, R.; Puchler, A.
 Mol. Gen. Genet. 216, 293-302, 1989
 A:Title: The Rhizobium meliloti fdxN gene encoding a ferredoxin-like protein is neces
 A:Reference number: JE0034; MUID:89313667
 A:Accession: JE0034
 A:Molecule type: DNA
 A:Residues: 1-64 <KLU>
 A:Cross-references: GB:X52662; NID:9288340; PIDN:CAA36890.1; PID:9288342
 A:Experimental source: strain 2011
 C:Genetics:
 A:Gene: fdxN
 C:Superfamily: ferredoxin 2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology
 C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein
 F:2-64/Product: ferredoxin 2[4Fe-4S] I #status predicted <MAT>
 F:3-63/Domain: ferredoxin 2[4Fe-4S] homology <FER>
 F:10/13/16/35/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
 F:20/39/42/51/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 13.3%; Score 48; DB 1; Length 64;
 Best Local Similarity 26.7%; Pred. No. 1.5e+02;
 Matches 12; Conservative 3; Mismatches 14; Indels 16; Gaps 1;

OY 21 GKIDCPNCPENGICLPGLAVGSFSCPCPDGFTDPNCSVVEV 65
 DB 29 GEXYVDPYTKCN-----ECKGSGFTQGCASVCY 57

RESULT 15
 A25775
 metallothionein A - sea urchin (Strongylocentrotus purpuratus)
 C:Species: Strongylocentrotus purpuratus (purple urchin)
 C:Date: 22-Jul-1987 #sequence_revision 22-Jul-1987 #text_change 20-Aug-1999
 A:Accession: A25775
 R:Nemer, M.; Wilkinson, D.G.; Travaaglini, E.C.; Sternberg, E.J.; Butte, T.R.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4592-4594, 1985
 A:Title: Sea urchin metallothionein sequence: key to an evolutionary diversity.
 A:Reference number: A25775; MUID:85270452
 A:Accession: A25775
 A:Molecule type: mRNA
 A:Residues: 1-64 <NEK>

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:33:10 : Search time 9.47 Seconds
(Without alignments)
235.122 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Perfect score: 362
Sequence: 1 MKRSVAWMLVGLSLGVPQF.....FSCBCEPDGFTDPCSSVVEY 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 5422

Minimum DB seq length: 0

Maximum DB seq length: 65

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69.5	19.2	53	1 EGF_PIG	Q00968 sus scrofa
2	56.5	15.6	1	MT_PERFL	P52725 perfla
3	56.5	15.1	60	1 MTCU_HELPO	P55447 helix pomat
4	54.5	15.1	64	1 MT_STENE	P55953 stenechilus
5	53	14.6	62	1 MT1_CANGA	P15113 candida gla
6	51.5	14.2	61	1 MT1_PAT	P02803 rattus norv
7	50.5	14.0	60	1 MT_PLEPL	P02716 pleuronecte
8	50.5	14.0	60	1 MT_PSEAM	P55945 pseudopleur
9	49.5	13.7	64	1 MTA_STRUV	P04734 strongyloce
10	49.5	13.5	64	1 MTA_PRUV	O48951 prunus aviu
11	49	13.5	65	1 FERN_AZOC	P06123 azotobacter
12	48.5	13.4	60	1 MT_GADMO	P51902 gadus morhu
13	48.5	13.4	60	1 MT_PAGBE	O92145 pagothenia
14	48.5	13.4	62	1 MTA_HUMAN	P47944 homo sapien
15	48	13.3	64	1 FERN_RHIME	P12712 rhizobium m
16	47.5	13.1	60	1 MT_CHAAC	P52724 chaeocephala
17	47	13.0	62	1 MTA_MOUSE	P47945 mus musculu
18	46.5	12.8	56	1 MTA_MOUSE	P56682 apis mellif
19	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
20	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
21	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
22	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
23	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
24	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
25	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
26	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
27	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
28	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
29	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
30	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
31	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
32	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch
33	46.5	12.8	60	1 MTA_MOUSE	O99899 dicentrarch

34	44.5	12.3	48	1 PIGU_HIRNT	P81499 hirudo nrip
35	44.5	12.3	60	1 MTA_HIRNT	P81499 hirudo nrip
36	44.5	12.3	60	1 MTA_HIRNT	P81499 hirudo nrip
37	44	12.2	38	1 BD01_BOVIN	P46159 bos taurus
38	44	12.2	58	1 FERG_METJA	O58041 methanococc
39	44	12.2	60	1 MT_NOEBA	P25128 noemacheilus
40	44	12.2	61	1 MTA_HORSE	P02800 equus caball
41	44	12.2	62	1 MT_XENLA	O05890 xenopus lae
42	43.5	12.0	42	1 BD13_BOVIN	P46171 bos taurus
43	43.5	12.0	61	1 MTA_MOUSE	P02802 mus musculu
44	43.5	12.0	64	1 BD01_PIG	O62697 sus scrofa
45	43	11.9	39	1 DECO_MACDE	P17350 macrobellia

ALIGNMENTS

RESULT	ID	EGF_PIG	STANDARD	PRT	53 AA
AC	Q00968				
DF	01-OCF-1996 (Rel. 34, last sequence update)				
DF	01-OCF-1996 (Rel. 34, last sequence update)				
DF	01-OCF-1996 (Rel. 34, last sequence update)				
DE	EPIDERMAL GROWTH FACTOR (EGF) (FRAGMENT)				
GN	EGF				
OS	Sus scrofa (Pig)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.				
OX	NCBI_taxonomy:9823;				
RM	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE-KIDNEY:				
RC	MEDLINE-91197366; PubMed-2015058;				
RA	Pascall J.C., Jones D.S.C., Odel S.M., Clements J.M., Hunter M.,				
RA	Fallon T., Edwards M., Brown K.D.				
RT	Cloning and characterization of a gene encoding pig epidermal growth				
RT	factor.				
RT	J. Mol. Endocrinol. 6:63-70(1991).				
CC	-1- FUNCTION: THE GROWTH FACTOR STIMULATES THE GROWTH OF VARIOUS				
CC	EPIDERMAL AND EPITHELIAL TISSUES IN VIVO AND IN VITRO AND OF SOME				
CC	FIBROBLASTS IN CELL CULTURE.				
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; X59516; CAA42102.1; -				
DR	HSSP; P01132; IEPH.				
DR	InterPro; IPR000561; -				
DR	Pfam; PF00008; EGF_1.				
DR	PROSITE; PS00022; EGF_1; 1.				
DR	PROSITE; PS0186; EGF_2; 1.				
KW	EGF-like domain; growth factor.				
FT	NON_TER	1			
FT	PEPTIDE	1			
FT	DOMAIN	2			
FT	DISULFID	6			
FT	DISULFID	14			
FT	DISULFID	33			
FT	NON_TER	53			
SEQUENCE	53 AA; 6149 MW; 74F615BA05774D4 CRC64;				

Query Match 19.2%; Score 69.5; DB 1; Length 53;
Best Local Similarity 35.1%; Pred. No. 0.33; 14; Indels 3; Gaps 1;
Matches 13; Conservative 7; Mismatches 14; Indels 3; Gaps 1;


```

AC P55953;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE METALLOTHIONEIN (MT).
OS Streptococcus neumannii (Antarctic sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Echinidae;
OC Streptococcus.
OX NCBI_TaxID=53479;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RX MEDLINE=97319675; PubMed=9176568;
RA Scudiero R., Capasso C., Carinae V., Riggio M., Capasso A.,
RA Carameella M., Filosa S., Partis E.;
RT "PCR amplification and cloning of metallothionein complementary DNAs
RT in temperate and Antarctic sea urchin characterized by a large
RL difference in egg metallothionein content.";
RL Cell. Mol. Life Sci. 53:472-477(1997).
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -1- SIMILARITY: BELONGS TO FAMILY 4 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y08621; CAA69912.1;
DR HSSP; P04355; 4MT2.
DR InterPro: IPR001396;
DR PRINTS; PR00873; MTECHINOIDEA
KW Metal-binding; Metal-thiolate cluster; Chelation.
SQ SEQUENCE 64 AA; 6428 MM; 326558A25CA0BA CRC64;

```

```

Query Match 15.1%; Score 54.5; DB 1; Length 64;
Best Local Similarity 30.0%; Pred. No. 15;
Matches 15; Conservative 5; Mismatches 15; Indels 15; Gaps 3;

```

```

OY 22 KGDCIPNCPENGICLPGIANGV---SPSCPCPG-----FDPNCS 60
DB 18 KGRKC-----CTTGECCKDGTGCGCKTNAACKADGCGSGCSCTRGNA 63

```

```

RESULT 5
MTL_CANGA STANDARD; PRT; 62 AA.
AC P15113;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN-I.
GN MT-1.
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Mitothioric Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90062075; PubMed=2584191;
RA Mehra R.K., Garey J.R., Butt T.R., Gray W.R., Winge D.R.;
RT "Candida glabrata metallothioneins. Cloning and sequence of the genes
RT and characterization of proteins.";
RL J. Biol. Chem. 264:19747-19753(1989).
RN [2]
RP SEQUENCE OF 1-16.
RX MEDLINE=89057829; PubMed=3194392;
RA Mehra R.K., Tarbet B.E., Gray W.R., Winge D.R.;

```

```

RT "Metal-specific synthesis of two metallothioneins and gamma-glutamyl
RT peptides in Candida glabrata.";
RT Proc. Natl. Acad. Sci. U.S.A. 85:8815-8819(1988).
CC -1- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -1- INDUCTION: BOTH MT-I AND MT-II GENES ARE REGULATED BY COPPER ION
CC IN A CONCENTRATION-DEPENDENT FASHION, AND BOTH ARE INDUCIBLE BY
CC SILVER BUT NOT BY CADMIUM SALTS.
CC -1- MISCELLANEOUS: MT-1 MIGHT BIND APPROXIMATELY 11-12 MOL EQ OF
CC CU(I).
CC -1- SIMILARITY: BELONGS TO FAMILY 9 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05133; AAA5272.1;
DR PIR; A31252; A31252.
DR PIR; A34484; A34484.
DR HSSP; P04355; 4MT2.
KW Metal-binding; Metal-thiolate cluster; Chelation; Copper; Repeat.
FT INIT_MET 0
FT REPEAT 22
FT REPEAT 29
FT REPEAT 55
SQ SEQUENCE 62 AA; 6243 MM; 80D768C06C44F7A1 CRC64;

```

```

Query Match 14.6%; Score 53; DB 1; Length 62;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 48 CCRPDGFTDPNCS 61
DB 4 CCRPDGFTDPNCS 17

```

```

RESULT 6
MTL_RAT STANDARD; PRT; 61 AA.
AC P02803;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE METALLOTHIONEIN-I (MT-1).
GN MT1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87064300; PubMed=3023830;
RA Andersen R.D., Birren B.W., Taplitz S.J., Herschman H.R.;
RT "Rat metallothionein-1 structural gene and three pseudogenes, one of
RT which contains 5'-regulatory sequences.";
RL Mol. Cell. Biol. 6:302-314(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=83209120; PubMed=6687866;
RA Andersen R.D., Birren B.W., Ganz T., Piletz J.E., Herschman H.R.;
RT "Molecular cloning of the rat metallothionein I (MT-1) mRNA
RL DNA 2:15-22(1983).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88029906; PubMed=2959527;
RA Andersen R.D., Taplitz S.J., Birren B.W., Bristol G., Herschman H.R.;
RT "Rat metallothionein multigene family.";

```

RL Experiencedia Suppl. 52:373-384(1987).
 RN [4]
 RP MEDLINE-84289593; PubMed-6470004;
 RX Winge D.R., Nelson K.B., Zeikus R.D., Gray W.R.;
 RA "Structural characterization of the isoforms of neonatal and adult
 RT rat liver metallothionein".
 RL J. Biol. Chem. 259:11419-11425(1984).
 RN [5]
 RP MEDLINE-83073694; PubMed-6184083;
 RX Mercer J.F.B., Hudson P.J.;
 RA "Cloning of metallothionein cDNA from neonatal rat liver.";
 RL Biosci. Rep. 2:761-768(1982).
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
 CC GLUCOCORTICOID.
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS;
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J00750; AAA41590.1; -;
 DR EMBL: M1794; AAA41541.1; -;
 DR EMBL: M24327; AAA41589.1; -;
 DR PIR: A03281; SMRT1.
 DR PIR: A61561; A61561.
 DR HSSP: P02795; 1MHU.
 DR InterPro: IPR000006; -;
 DR InterPro: IPR003019; -;
 DR Pfam: PF00131; metalthio.1.
 DR PRINTS: PR00860; MVTERTERRATE.
 DR PROSITE: PS00203; METALLOTHIONEIN_VRT.1.
 KM Metal-binding; Metal-thiolate cluster; Chelation; Zinc; Acetylation.
 FT MOD_RES 1 29
 FT DOMAIN 1 30 61
 FT METAL 5 5
 FT METAL 7 7
 FT METAL 13 13
 FT METAL 15 15
 FT METAL 19 19
 FT METAL 21 21
 FT METAL 24 24
 FT METAL 26 26
 FT METAL 29 29
 FT METAL 33 33
 FT METAL 34 34
 FT METAL 36 36
 FT METAL 37 37
 FT METAL 41 41
 FT METAL 44 44
 FT METAL 48 48
 FT METAL 50 50
 FT METAL 57 57
 FT METAL 59 59
 FT METAL 60 60
 SQ SEQUENCE 61 AA; 6006 MW; 446083720B9ABC18 CRC64;

Query Match 14.2%; Score 51.5; DB 1; Length 61;
 Best Local Similarity 40.0%; Pred. No. 30;

Matches 10; Conservative 3; Mismatches 11; Indels 1; Gaps 1;
 QY 27 DPN-PCENGICLPGLAWSEFCSEC 50
 DB 2 DPNCGSTGSGCTCSCSGCKNKC 26

RESULT 7
 MT_PLEPL STANDARD; PRT; 60 AA.
 ID MT_PLEPL
 AC P07216;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE METALLOTHIONEIN (MT).
 GN MT.
 OS Pleurocetes platessa (Plaice).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectidae; Pleuronectidae; Pleuronectes.
 OX NCBI_TaxID=8262;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Leaver M.J., George S.G.;
 RN Submitted (NOV-1990) to the EMBL/Genbank/DBJ databases.
 RL [2]
 RP PRELIMINARY PARTIAL SEQUENCE.
 RT Overnell J., Berger C., Wilson K.J.;
 RT "Partial amino acid sequence of metallothionein from the plaice
 (Pleuronectes platessa)".
 RL Biochem. Soc. Trans. 9:217-218(1981).
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS;
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X56743; GAA40067.1; -;
 DR PIR: A05076; A05076.
 DR PIR: S30567; S30567.
 DR HSSP: F18055; 1MRB.
 DR InterPro: IPR000006; -;
 DR InterPro: IPR003019; -;
 DR Pfam: PF00131; metalthio.1.
 DR PRINTS: PR00860; MVTERTERRATE.
 DR PROSITE: PS00203; METALLOTHIONEIN_VRT.1.
 KM Metal-binding; Metal-thiolate cluster; Chelation; Copper; Acetylation.
 FT MOD_RES 1 1
 FT DOMAIN 1 1
 FT METAL 29 26
 FT METAL 4 4
 FT METAL 6 6
 FT METAL 12 12
 FT METAL 14 14
 FT METAL 18 18
 FT METAL 20 20
 FT METAL 23 23
 FT METAL 25 25
 FT METAL 28 28
 FT METAL 32 32
 SQ SEQUENCE 32 AA; 3206 MW; 446083720B9ABC18 CRC64;


```

FT METAL 33 33 CLUSTER A.
FT METAL 35 35 CLUSTER A.
FT METAL 36 36 CLUSTER A.
FT METAL 40 40 CLUSTER A.
FT METAL 43 43 CLUSTER A.
FT METAL 47 47 CLUSTER A.
FT METAL 49 49 CLUSTER A.
FT METAL 54 54 CLUSTER A.
FT METAL 58 58 CLUSTER A.
FT METAL 59 59 CLUSTER A.
SQ SEQUENCE 60 AA; 6029 MM; 67B046EA4C7C470C CRC64;

Query Match 14.0%; Score 50.5; DB 1; Length 60;
Best Local Similarity 30.6%; Pred. No. 37;
Matches 15; Conservative 3; Mismatches 12; Indels 19; Gaps 3;

OY 27 DPNPCNGGICLPGLAVGSFSC-----CPDGFDPNCS 61
    || | | | | | | | | | | | | | | | | | | | |
DB 2 DPCCKSKGTGTCNGG--GSCTCKNCSCTCNKSCPCPCSGC--PKCAS 45

RESULT 8
MT_PSEAM STANDARD; PRT; 60 AA.
ID MT_PSEAM
AC P53945;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE METALLOTHIONEIN (MT).
GN MT.
OS Pseudopleuronectes americanus (Winter flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorphi; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Pseudopleuronectes.
OX NCBI_TaxID=8265;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-25.
RC TISSUE=Liver.
RA Chan K.-M., Davidson W.S., Hew C.-L., Fletcher G.L.;
RT "Molecular cloning of metallothionein cDNA and analysis of
RT metallothionein gene expression in winter flounder tissues."
RL Can. J. Zool. 67:2520-2527(1989).
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLEATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
DR HSSP: P18055; IMRB.
DR InterPro: IPR000006; -.
DR InterPro: IPR003019; -.
DR Pfam: PF00131; metalthio_1.
DR PRINTS: PR00860; MTVEITERATE.
DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Chelation; Copper; Acetylation.
FT MOD_RES 1 28 ACETYLATION.
FT DOMAIN 1 28 BETA.
FT METAL 29 60 ALPHA.
FT METAL 4 4 CLUSTER B (BY SIMILARITY).
FT METAL 6 6 CLUSTER B (BY SIMILARITY).
FT METAL 12 12 CLUSTER B (BY SIMILARITY).
FT METAL 14 14 CLUSTER B (BY SIMILARITY).
FT METAL 18 18 CLUSTER B (BY SIMILARITY).
FT METAL 20 20 CLUSTER B (BY SIMILARITY).
FT METAL 23 23 CLUSTER B (BY SIMILARITY).
FT METAL 25 25 CLUSTER B (BY SIMILARITY).
FT METAL 26 26 CLUSTER B (BY SIMILARITY).
FT METAL 32 32 CLUSTER A (BY SIMILARITY).
FT METAL 33 33 CLUSTER A (BY SIMILARITY).

```

```

FT METAL 35 35 CLUSTER A (BY SIMILARITY).
FT METAL 36 36 CLUSTER A (BY SIMILARITY).
FT METAL 40 40 CLUSTER A (BY SIMILARITY).
FT METAL 43 43 CLUSTER A (BY SIMILARITY).
FT METAL 47 47 CLUSTER A (BY SIMILARITY).
FT METAL 49 49 CLUSTER A (BY SIMILARITY).
FT METAL 54 54 CLUSTER A (BY SIMILARITY).
FT METAL 58 58 CLUSTER A (BY SIMILARITY).
FT METAL 59 59 CLUSTER A (BY SIMILARITY).
SQ SEQUENCE 60 AA; 6043 MM; 67B046E5C7C470C CRC64;

Query Match 14.0%; Score 50.5; DB 1; Length 60;
Best Local Similarity 30.6%; Pred. No. 37;
Matches 15; Conservative 3; Mismatches 12; Indels 19; Gaps 3;

OY 27 DPNPCNGGICLPGLAVGSFSC-----CPDGFDPNCS 61
    || | | | | | | | | | | | | | | | | | | | |
DB 2 DPCCKSKGTGTCNGG--GSCTCKNCSCTCNKSCPCPCSGC--PKCAS 45

RESULT 9
MTA_STRPU STANDARD; PRT; 64 AA.
ID MTA_STRPU
AC P04734;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE METALLOTHIONEIN-A (MTA).
GN MTA.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinoidea; Echinodermata; Echinozoa; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9006495; Pubmed=2586524;
RA Harlow P., Watkins E., Thornton R.D., Nemer M.;
RT "Structure of an ecdoderminally expressed sea urchin metallothionein
RT gene and characterization of its metal-responsive region."
RL Mol. Cell. Biol. 9:5445-5455(1989).
CC [2]
CC SEQUENCE FROM N.A.
CC MEDLINE=85270452; Pubmed=3860837;
CC Nemer M., Wilkinson D.G., Travaglini E.C., Sternberg E.J., Butt T.R.;
CC "Sea urchin metallothionein sequence: key to an evolutionary
CC diversity."
CC Proc. Natl. Acad. Sci. U.S.A. 82:4992-4994(1985).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=99370086; Pubmed=10438629;
RA Riek R., Prenchet B., Wang Y., Mackay E.A., Wilder G., Guntert P.,
RA Ily A., Kaegi J.H.R., Wuthrich K.;
RT "NMR structure of the sea urchin (Strongylocentrotus purpuratus)
RT metallothionein MTA."
CC J. Mol. Biol. 291:417-428(1999).
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -1- INDUCTION: BELONGS TO FAMILY 4 IN METALLOTHIONEIN SUPERFAMILY.
CC -1- SIMILARITY: BELONGS TO FAMILY 4 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M30606; AAA30067.1; -.
DR EMBL: K02464; AAA30061.1; -.
DR PIR: A3825; A3825.

```

DR PDB: 10JK: 31-AUG-99.
 DR PDB: 10JK: 31-AUG-99.
 DR InterPro: IPR001396: -
 DR PRINTS: PR00873: METECHINIDRA.
 KM Metal-binding; Metal-chelate cluster; Chelation; 3D-structure.
 FT CONFLICT 10 K -> T (IN REF. 2).
 FT CONFLICT 24 K -> V (IN REF. 2).
 SQ SEQUENCE 64 AA: 6444 MW: 66FDAA/4499IE3E6 CRC64:

Query Match 13.7%; Score 49.5; DB 1; Length 64;
 Best Local Similarity 28.6%; Pred. No. 50;
 Matches 14; Conservative 8; Mismatches 16; Indels 11; Gaps 4;

QY 20 FGKGIQDPNCPENGICILPGIANGSFSCPDG-----FTDPNCS 60
 DB 18 FQD-DCKTGCCKGDKGCC-GICTNA-ACKKANGCKGSGSCTEGNC 63

RESULT 10
 ID MT3_PROAV STANDARD: PRT: 64 AA.
 AC 048951.
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE METALLOTHIONEIN-LIKE PROTEIN 1.
 GN MT1.
 OS Prunus avium (Cherry).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1;
 OC Rosales; Rosaceae; Prunus.
 OX NCBI_TaxID=42229;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV SUMMITT; TISSUE-Fruit;
 RA Maerzma P.A., Wu Z., Wilson S.M.;
 RT A fruit-related metallothionein-like cDNA clone from sweet cherry
 RL corresponds to fruit genes from diverse species.";
 RL (in) Plant Gene Register PCR98-015.
 CC -1- FUNCTION: METALLOTHIONEIN HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 15 IN METALLOTHIONEIN SUPERFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF028013: AAC04612.1: -
 KM Metal-binding; Metal-chelate cluster; Chelation.
 SQ SEQUENCE 64 AA: 6743 MW: 12D8602A22ED109C CRC64:

Query Match 13.5%; Score 49; DB 1; Length 64;
 Best Local Similarity 38.5%; Pred. No. 57;
 Matches 10; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY 25 ICDPNCENGICILPGIANGSFSCSC 50
 DB 39 IMDAPAEANGCKGCGSCACVCKC 64

RESULT 11
 ID FERN_AZOGH STANDARD: PRT: 65 AA.
 AC P06123.
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE FERREDOXIN-LIKE PROTEIN IN VNF REGION.
 OS Azotobacter chroococcum mcd 1.
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Azotobacter.
 OX NCBI_TaxID=355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-90356423; PubMed-2388847;
 RA Fallik E., Robson R.L.;
 RT "Completed sequence of the region encoding the structural genes for
 RT the vanadium nitrogenase of Azotobacter chroococcum.";
 RN Nucleic Acids Res. 18:4616-4616(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Robson R., Woodley P., Jones R.;
 RT "Second gene (nifH*) coding for a nitrogenase iron protein in
 RT Azotobacter chroococcum is adjacent to a gene coding for a ferredoxin-
 RT like protein.";
 RL EMBO J. 5:1159-1163(1986).
 CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN.
 CC STRONGEST SIMILARITY TO OTHER NITROGEN-FIXING FERREDOXINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X51756: CAA36056.1: -
 DR EMBL: X03916: CAA27554.1: -
 DR PTR: B25103; B25103.
 DR PTR: S14693; S14693.
 DR HSSP: P00193; IFDX.
 DR InterPro: IPR001450: -
 DR POSITIVE: PS00198; 4FEAS-FERREDOXIN; 1.
 DR PROSITE: PS00198; 4FEAS-FERREDOXIN; 1.
 KM Electron transport; Iron-sulfur; 4Fe-4S; Duplication;
 KW Nitrogen fixation.
 FT METAL 10 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 13 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 16 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 16 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 FT METAL 20 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 39 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 42 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 50 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 54 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
 SQ SEQUENCE 65 AA: 6804 MW: 271FA6FC58B2FAB3 CRC64:

Query Match 13.5%; Score 49; DB 1; Length 65;
 Best Local Similarity 39.1%; Pred. No. 58;
 Matches 18; Conservative 3; Mismatches 19; Indels 6; Gaps 5;

QY 23 GDICDPNCPENGICILPGIANGSFSC-ECPPDFTDPNCSVVEV 65
 DB 14 GD-CRP-VCPGTSTVLDGIVYIDADSNCWCAD-LGEPRCIGVCPV 56

RESULT 12
 ID MT_GADMO STANDARD: PRT: 60 AA.
 AC P51902: 092078;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE METALLOTHIONEIN (MT).
 GN MT.
 OS Gadus morhua (Atlantic cod).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadidae;
OC Gadus;
OX NCBI_TaxID=8049;
RN (1)
RN SEQUENCE FROM N.A.
RX MEDLINE=95187262; PubMed=7881512;
RA McNamara P.T., Buckley L.J.;
RT "Identification and characterization of metallothionein cDNA from
RT mRNA transcripts induced by starvation in Atlantic cod (Gadus
RT morhua).";
RL Mol. Mar. Biol. Biotechnol. 3:252-260(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP Tissue=Brain;
RA Kille P., Olsson P.E.;
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U08105; AAA4418.1; -
DR EMBL: X97269; CAA65924.1; -
DR HSSP: P18055; 2MRB.
DR InterPro: IPR000006; -
DR Pfam: PF00131; metalthio.1.
DR PRINTS: PR00860; MTEVERBRATE.
DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.
KM Metal-binding; Metal-thiolate cluster; Chelation.
FT DOMAIN 1 28 BETA.
FT METAL 29 60 ALPHA.
FT METAL 4 4 CLUSTER B.
FT METAL 6 6 CLUSTER B.
FT METAL 12 12 CLUSTER B.
FT METAL 14 14 CLUSTER B.
FT METAL 18 18 CLUSTER B.
FT METAL 20 20 CLUSTER B.
FT METAL 23 23 CLUSTER B.
FT METAL 25 25 CLUSTER B.
FT METAL 28 28 CLUSTER B.
FT METAL 32 32 CLUSTER A.
FT METAL 33 33 CLUSTER A.
FT METAL 35 35 CLUSTER A.
FT METAL 36 36 CLUSTER A.
FT METAL 40 40 CLUSTER A.
FT METAL 43 43 CLUSTER A.
FT METAL 47 47 CLUSTER A.
FT METAL 49 49 CLUSTER A.
FT METAL 54 54 CLUSTER A.
FT METAL 58 58 CLUSTER A.
FT METAL 59 59 CLUSTER A.
FT METAL 5 5 D -> E (IN REF. 2).
FT CONFLICT 7 7 A -> S (IN REF. 2).
FT CONFLICT 11 11 T -> S (IN REF. 2).
FT CONFLICT 48 48 A -> V (IN REF. 2).
SQ SEQUENCE 60 AA; 6106 MW; B860D692009F3969 CRC4;

Matches 11; Conservative 3; Mismatches 17; Indels 3; Gaps 1;
QY 26 CDPNCGNGICLPIGLANGSESCPCDGFDPNC 59
DB 28 CKSCCE--CCPSGCKSCAGCACKDKTCDTNC 58
RESULT 13
MT PAGE
ID MT_PAGE STANDARD; PRT; 60 AA.
AC 092145;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE METALLOTHIONEIN (MT).
GN MT.
OS Pagodienia bernacchi (Pmerald rockcod) (Trematomus bernacchi), and
OS Chionodirco rastrospinosus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
OC Notothenioidae; Nototheniidae; Trematomus.
OX NCBI_TaxID=40690, 34790;
RN [1]
RN SEQUENCE FROM N.A.
RP Tissue=Liver;
RA Scudiero R.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z72485; CAA96566.1; -
DR EMBL: Z72484; CAA96565.1; -
DR HSSP: P18055; 1MRB.
DR InterPro: IPR000006; -
DR Pfam: PF00131; metalthio.1.
DR PRINTS: PR00860; MTEVERBRATE.
DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.
KM Metal-binding; Metal-thiolate cluster; Chelation.
FT DOMAIN 1 28 BETA.
FT METAL 29 60 ALPHA.
FT METAL 4 4 CLUSTER B.
FT METAL 6 6 CLUSTER B.
FT METAL 12 12 CLUSTER B.
FT METAL 14 14 CLUSTER B.
FT METAL 18 18 CLUSTER B.
FT METAL 20 20 CLUSTER B.
FT METAL 23 23 CLUSTER B.
FT METAL 25 25 CLUSTER B.
FT METAL 28 28 CLUSTER B.
FT METAL 32 32 CLUSTER A.
FT METAL 33 33 CLUSTER A.
FT METAL 35 35 CLUSTER A.
FT METAL 36 36 CLUSTER A.
FT METAL 40 40 CLUSTER A.
FT METAL 43 43 CLUSTER A.
FT METAL 47 47 CLUSTER A.
FT METAL 49 49 CLUSTER A.

FT METAL 54 54 CLUSTER A.
 FT METAL 58 58 CLUSTER A.
 FT METAL 59 59 CLUSTER A.
 SQ SEQUENCE 60 AA: 5992 MW: E866F4AD61C424A CRC64.

Query Match 13.4%; Score 48.5; DB 1; Length 60;
 Best Local Similarity 30.6%; Pred. No. 61;
 Matches 13; Conservative 2; Mismatches 13; Indels 19; Gaps 3;

OY 27 DPNCENGICLPGLAVGFSFCE-----CPDGFDPNCSS 61
 Db 2 DPECVMSGICMCG-----DNCRC-----TTGNCMT 29

RESULT 14

MT4_HUMAN STANDARD; PRT; 62 AA.

AC P47944;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE METALOTHIONEIN-IV (MT-IV).

GN MT4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

NCBI_TaxId=9606;

RA [1]

SEQUENCE FROM N.A.

RA MEDLINE=9427179; PubMed=8003488;

RA Kelle E.J., Findey S.D., Erickson J.C., Froelich G.J.,

RA Kelly E.J., Zambrowicz B.P., Palmer R.D.,

RT Induction of a new metallothionein isoform (MT-IV) occurs during

RT differentiation of stratified squamous epithelia."

RL Biochemistry 33:7250-7259(1994).

-1- FUNCTION: SEEMS TO BIND ZINC AND COPPER. COULD PLAY A SPECIAL ROLE

IN REGULATING ZINC METABOLISM DURING THE DIFFERENTIATION OF

STRATIFIED EPITHELIA.

-1- SIMILARITY: BELONGS TO FAMILY 1 IN METALOTHIONEIN SUPERFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U07807; AAA20232.1; -

DR HSP: P18055; 1MRB

DR InterPro: IPR000006; -

DR InterPro: IPR003019; -

DR Pfam: PF00131; metalthio; 1.

DR PRIS: PR00860; MYTEREPRATE.

DR PROSITE: PS00203; METALOTHIONEIN_VRT; 1.

KW Metal-binding; Metal-thiolate cluster; Zinc; Copper.

FT METAL 6 6 CLUSTER B (BY SIMILARITY).
 FT METAL 8 8 CLUSTER B (BY SIMILARITY).
 FT METAL 14 14 CLUSTER B (BY SIMILARITY).
 FT METAL 16 16 CLUSTER B (BY SIMILARITY).
 FT METAL 20 20 CLUSTER B (BY SIMILARITY).
 FT METAL 22 22 CLUSTER B (BY SIMILARITY).
 FT METAL 25 25 CLUSTER B (BY SIMILARITY).
 FT METAL 27 27 CLUSTER B (BY SIMILARITY).
 FT METAL 30 30 CLUSTER B (BY SIMILARITY).
 FT METAL 34 34 CLUSTER A (BY SIMILARITY).
 FT METAL 35 35 CLUSTER A (BY SIMILARITY).
 FT METAL 37 37 CLUSTER A (BY SIMILARITY).
 FT METAL 38 38 CLUSTER A (BY SIMILARITY).
 FT METAL 42 42 CLUSTER A (BY SIMILARITY).
 FT METAL 45 45 CLUSTER A (BY SIMILARITY).
 FT METAL 49 49 CLUSTER A (BY SIMILARITY).

FT METAL 51 51 CLUSTER A (BY SIMILARITY).
 FT METAL 58 58 CLUSTER A (BY SIMILARITY).
 FT METAL 60 60 CLUSTER A (BY SIMILARITY).
 FT METAL 61 61 CLUSTER A (BY SIMILARITY).
 SQ SEQUENCE 62 AA: 6419 MW: 36157CB1A7BF28CC CRC64.

Query Match 13.4%; Score 48.5; DB 1; Length 62;
 Best Local Similarity 35.1%; Pred. No. 63;
 Matches 13; Conservative 5; Mismatches 8; Indels 11; Gaps 3;

OY 27 DPNCENGICLPGLAVGFSFCECPDGFDPNCSS 61
 Db 2 DPECVMSGICMCG-----DNCRC-----TTGNCMT 29

RESULT 15

FERN_RHIME STANDARD; PRT; 64 AA.

AC P12712;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE FERREDOXIN-LIKE PROTEIN IN NIF REGION.

GN FDHN.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Plasmid psym (megaplasmid 1).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

NCBI_TaxId=382;

RA [1]

SEQUENCE FROM N.A.

RA MEDLINE=8831954; PubMed=2842320;

RA Mulligan M.E., Bulken W.J., Haselkorn R.;

RT "Bacterial-type ferredoxin genes in the nitrogen fixation regions of

RT the cyanobacterium Anabaena sp. strain PCC 7120 and Rhizobium

RT meliloti."

RL J. Bacteriol. 170:4406-4410(1988).

RL [2]

SEQUENCE FROM N.A.

RC STRAIN=RCR2011 / SU47;

RA MEDLINE=8931367; PubMed=2747618;

RA Klipp W., Reilander H., Schlueter A., Krey R., Puchler A.;

RT "The Rhizobium meliloti fdxN gene encoding a ferredoxin-like protein

RT nifB."

RL Mol. Gen. Genet. 216:293-302(1989).

RL [3]

3D-STRUCTURE MODELING OF IRON-SULFUR CLUSTERS.

RA MEDLINE=92293135; PubMed=163075;

RA Masopohl B., Kutsche M., Riedel K.-U., Schmel M., Klipp W.;

RT "Functional analysis of the cysteine motifs in the ferredoxin-like

RT protein fdxN of Rhizobium meliloti involved in symbiotic nitrogen

RT fixation."

RL Mol. Gen. Genet. 233:33-41(1992).

-1- CORRELATOR: BINDS TWO 4FE-4S CLUSTERS (PROBABLE).

-1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.

STRONGEST SIMILARITY TO OTHER NITROGEN-FIXING FERREDOXINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M21841; AAA26268.1; -

DR EMBL: X52662; CAA36890.1; -

DR PIR: G32361; G32361

DR PIR: JF0034; JF0034.

DR HSP: P00198; 1FCA.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:32:54 ; Search time 19.15 Seconds
(without alignments)
397.834 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97
Sequence: 1 MKRSVAVMLVGLSGVPOF.....FSCPCPDGTPDNCSSVEV 65
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 35489

Minimum DB seq length: 0

Maximum DB seq length: 65

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL.15:*
2: SP_archaea:*
3: SP_bacteria:*
4: SP_fungi:*
5: SP_human:*
6: SP_invertebrate:*
7: SP_mammal:*
8: SP_ornithine:*
9: SP_phage:*
10: SP_plant:*
11: SP_ricent:*
12: SP_unclassified:*
13: SP_vertebrate:*
14: SP_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77.5	21.4	43	5	Q9TX98
2	56.2	17.1	56	11	Q9QMB5
3	56.5	16.2	47	4	Q9E077
4	54.5	15.1	33	4	Q9S518
5	54.5	15.1	36	2	Q9E85
6	53.2	14.4	65	6	Q77682
7	50.5	14.0	48	13	Q9E855
8	30	13.8	63	10	Q23538
9	30	13.8	65	2	Q47784
10	30	13.7	62	6	Q9T015
11	48.5	13.7	62	6	Q9T015
12	48.5	13.7	64	2	Q9Z831
13	48.5	13.4	60	13	Q93609
14	48.5	13.4	60	13	Q13259
15	48.5	13.4	60	13	Q13259
16	48.5	13.4	60	13	Q13258
17	48.5	13.3	39	6	Q9M277
18	46.5	12.8	60	13	Q93593
19	46.5	12.8	60	13	Q91B50

20	46	12.7	39	5	Q9TX98	Q9TX98 caenorhabdi
21	45.5	12.6	63	4	Q9NNX2	Q9NNX2 homo sapien
22	45	12.4	50	4	Q9UBA5	Q9UBA5 homo sapien
23	45	12.4	61	6	Q18842	Q18842 balena mys
24	44.5	12.3	46	10	Q42259	Q42259 arabidopsis
25	44	12.2	44	11	Q60620	Q60620 mus musculu
26	44	12.2	57	5	Q9N9H2	Q9N9H2 ruditapes d
27	44	12.2	59	5	Q9N9H1	Q9N9H1 ruditapes d
28	44	12.2	61	6	P79376	P79376 sus scrofa
29	44	12.2	61	6	P79376	P79376 sus scrofa
30	43.5	12.0	48	5	Q9TXA0	Q9TXA0 caenorhabdi
31	43	11.9	48	6	P79380	P79380 cypripus ca
32	43	11.9	60	13	Q91910	Q91910 sus scrofa
33	43	11.9	61	6	P79379	P79379 sus scrofa
34	43	11.9	62	6	Q19156	Q19156 cercopthec
35	43	11.9	63	1	Q93653	Q93653 archaeoglob
36	43	11.9	63	2	Q53548	Q53548 mycobacteri
37	43	11.9	64	10	Q40049	Q40049 hordeum vul
38	43	11.9	65	1	Q9YB46	Q9YB46 aeropyrum p
39	43	11.9	65	6	Q18958	Q18958 bos taurus
40	42.5	11.7	42	6	Q18958	Q18958 bos taurus
41	42.5	11.7	45	10	Q81529	Q81529 mesembryant
42	42.5	11.7	48	6	P79377	P79377 sus scrofa
43	42.5	11.7	60	5	Q97413	Q97413 anopheles g
44	42.5	11.7	61	6	Q9XST5	Q9XST5 canis famli
45	42.5	11.7	61	6	Q9XST5	Q9XST5 canis famli

ALIGNMENTS

RESULT 1	PRELIMINARY;	PRT:	43 AA.
ID Q9TX98			
AC Q9TX98			
DT 01-MAY-2000 (TRENDArel. 13, Created)			
DR 01-MAY-2000 (TRENDArel. 13, Last sequence update)			
DE 01-JUN-2000 (TRENDArel. 14, Last annotation update)			
DE GIP-1 PROTEIN (FRAGMENT)			
CC Caenorhabditis elegans			
CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;			
CC Rhabditidae; Pelodermidae; Caenorhabditis.			
OX NCBI_TaxID:6219;			
PN [1]			
RP SEQUENCE FROM N.A.			
RY MEDLINE:93091267; PubMed:1457827;			
RA Kodorianni V., Wayne P.M., Kimble J.;			
RT "Molecular basis of loss-of-function mutations in the gip-1 gene of			
RL Caenorhabditis elegans"; 1991-1213(1992).			
DR NCBI_Biol. Citl 31199-1213(1992).			
DR INTERPRO: IPR000561;			
DR INTERPRO: IPR001438;			
DR PRINTS: PR00010; E09B10D.			
DR PROSITE: PS00022; G0F_1; UNKNOWN_1.			
DR PROSITE: PS01186; G0F_2; 1.			
SO SEQUENCE 43 AA; 4661 MW; 6C86E2421E3DA258 CRC64;			
Query Match	21.4%;	Score 77.5;	DB 5; Length 43;
Best Local Similarity	38.1%;	Pred. No. 0.016;	
Matches	16;	Conservative	5; Mismatches 18; Indels 3; Gaps 2;
18	PQFGKGDIDPCBNGSICLPGLAVGSFSCPCPDGTPDNC	59	
DB	4	PLFSGVNPQSDSPC-NNGICTP--FYGEFQICNNGTGGSTC	42
RESULT 2	PRELIMINARY;	PRT:	56 AA.
ID Q9QMB5			
AC Q9QMB5			
DT 01-MAY-2000 (TRENDArel. 13, Created)			
DR 01-MAY-2000 (TRENDArel. 13, Last sequence update)			

DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)
 DE GP 55-MAJOR EXT-GLOBULE MEMBRANE PROTEIN/MGF-E8 HOMOLOG (FRAGMENTS).
 OS Cavia (guinea pigs).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathii; Caviidae.
 ON NCBI_Taxid=10140;
 RN [1]
 RP SEQUENCE.
 RA Mather I.H., Banghart L.R., Lane W.S.;
 RL Biochem. Mol. Biol. Int. 29:545-554(1993).
 FT NON_TER 1 1
 FT NON_CONS 20 21
 FT NON_CONS 27 28
 FT NON_CONS 41 42
 FT NON_TER 56 56
 SQ SEQUENCE 56 AA; 6058 MW; 9861C0352393025D CRC64;

Query Match 17.1%; Score 62; DB 11; Length 56;
 Best Local Similarity 61.1%; Pred. No. 1.6; Mismatches 6; Indels 0; Gaps 0;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 23 GDIDPMPGNGSICLPG 40
 DB 3 GDFCDSSICLNGSTCLFG 20

RESULT 3
 ID Q9P0Z7 PRELIMINARY; PRT; 47 AA.
 AC Q9P0Z7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE LATENT TRANSFORMING GROWTH FACTOR BETA BINDING PROTEIN 3 (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Saharinen J., Koski C., Keski-Oja J.;
 RT "Cloning of human LTRP-3";
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: A0135961; AAF62353.1; -;
 FT NON_TER 1 1
 FT NON_TER 47 47
 SQ SEQUENCE 47 AA; 5086 MW; 57B1901D8B841B98 CRC64;

Query Match 16.2%; Score 58.5; DB 4; Length 47;
 Best Local Similarity 40.7%; Pred. No. 3.5; Mismatches 9; Indels 3; Gaps 1;
 Matches 11; Conservative 4; Mismatches 9; Indels 3; Gaps 1;
 QY 27 DPNCENGGICLPGIAY---GFSGSC 50
 DB 1 DVDECDPPACRPGRCVNLPESTRCC 27

RESULT 4
 ID Q95518 PRELIMINARY; PRT; 33 AA.
 AC Q95518;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE D1170K4.2 (EXON OF GENE 1170K4.2, EXACT EXON-STRUCTURE IN RELATION TO THE GENE UNKNOWN) (FRAGMENT).
 ON NCBI_Taxid=9606;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 ON NCBI_Taxid=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Blakey S.;
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: A022314; CA018442.2; -;
 DR HSSP: 007954; 1CR8.
 DR MEROPS: S01.308; -;
 DR INTERPRO: IPR001254; -;
 DR INTERPRO: IPR002172; -;
 DR PFAM: PF00057; 1d1_recept.a; 1.
 DR PFAM: PF00089; 1rypsin; 1-;
 DR PROSITE: PSS0068; 1d1RA.2; 1.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 3386 MW; EC7ACD3436B8C633 CRC64;

Query Match 15.1%; Score 54.5; DB 4; Length 33;
 Best Local Similarity 30.3%; Pred. No. 7.4; Mismatches 6; Indels 11; Gaps 2;
 Matches 10; Conservative 6; Mismatches 6; Indels 11; Gaps 2;
 QY 31 CENGGICLPGIAYSGFSC---ECPDGFTDPC 59
 DB 7 GSYNGLCVP-----ACDGVKDCPDGDERNC 32

RESULT 5
 ID Q9PAZ5 PRELIMINARY; PRT; 56 AA.
 AC Q9PAZ5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE HYPOTHETICAL PROTEIN XF2350.
 ON NCBI_Taxid=9606;
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 ON NCBI_Taxid=23711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-9A5C;
 RC MEDLINE-20365717; PubMed-10910347;
 RA Simpson A.J.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franco S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Gantier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Marques M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vicorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A., Paris A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Pasquero J.B.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pasquero J.B.,
 RA Quaggio B.R., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawaaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.2., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Seubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-157(2000).
 DR EMBL: AEO04045; AAF65149.1; -;
 KW Hypothetical protein.


```

RT *Characterization and heterologous expression of the genes encoding
RT enterococcal A production, immunity, and regulation in enterococcus
RT faecium DP0146."
RT Appl. Environ. Microbiol. 65:1506-1515(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-N15.
RA Uchiyama K., Lostekinkit C., Shioya S.;
RT "Characterization of bacteriocin produced by Enterococcus faecium N15
RT and cloning of the related genes."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: X94181; CAA63890.1; -.
DR EMBL: AF099088; AAD29132.1; -.
DR EMBL: AB038464; BAA92138.1; -.
DR HSSP: P34034; 2LEU.
DR INTERPRO: IPR002633; -.
DR PFAM: PF01721; Bacteriocin_11; 1.
DR PRODOM: PD004452; -. 1.
RW SIGNAL.
FT CHAIN 1 18
FT SIGNAL 19 65
SQ SEQUENCE 65 AA; 6887 MW; BCA9F966430B13AC CRC64;

Query Match 13.8%; Score 50; DB 2; Length 65;
Best Local Similarity 29.0%; Pred. No. 51;
Matches 9; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

OY 20 FGKGDICDPNPC-----ENGICICPLAVGSP 46
Db 26 YGNGYCTKNCCTVDMAKATTCIAGMSIGGF 56

RESULT 10
O9L568 PRELIMINARY; PRT; 65 AA.
AC O9L568;
DT 01-OCT-2000 (TREMREL. 15, Created)
DR 01-OCT-2000 (TREMREL. 15, Last sequence update)
DR 01-OCT-2000 (TREMREL. 15, Last annotation update)
DS CLASS IIA BACTERIOCIN ENTA.
EN ENTA.
OS Enterococcus faecium (Streptococcus faecium).
OC Bacteriis; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
CX NCBI_TaxID=1352;
NM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PC25;
RA Morovsky M., Pristas P., Javorsky P., Holo H., Nes I.F.;
RT "Characterization of Enterocoin BC25 bacteriocin and occurrence of entA
RT gene in rumen Gram-positive cocci."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF240561; AAF44686.1; -.
DR HSSP: P500198; 4FEAS_FERREROXIN; 1.
RW Iron-sulfur.
SQ SEQUENCE 65 AA; 6886 MW; ICBDED66430B08DA CRC64;

Query Match 13.8%; Score 50; DB 2; Length 65;
Best Local Similarity 29.0%; Pred. No. 51;
Matches 9; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

OY 20 FGKGDICDPNPC-----ENGICICPLAVGSP 46
Db 26 YGNGYCTKNCCTVDMAKATTCIAGMSIGGF 56

RESULT 11
O9T015 PRELIMINARY; PRT; 62 AA.
AC O9T015;
DT 01-MAY-2000 (TREMREL. 13, Created)
DR 01-MAY-2000 (TREMREL. 13, Last sequence update)
DR 01-OCT-2000 (TREMREL. 15, Last annotation update)

```

```

DE METALLOTHIONEIN-IV.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BEAGLE; TISSUE-TONGUE;
RA Kobayashi K., Morita T., Shimada A.;
RT "Molecular cloning and expression of metallothionein-IV."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB028041; BAA87326.1; -.
DR HSSP: P18055; 1MRB.
DR INTERPRO: IPR000006; -.
DR INTERPRO: IPR003019; -.
DR PFAM: PF00131; metalthio; 1.
DR PRINTS: PR00860; MTVERTERATE.
DR POSITE: PS00203; METALLOTHIONEIN_VRT; 1.
SQ SEQUENCE 62 AA; 6249 MW; 2858DAD17AA3DC9 CRC64;

Query Match 13.7%; Score 49.5; DB 6; Length 62;
Best Local Similarity 35.1%; Pred. No. 56;
Matches 13; Conservative 5; Mismatches 8; Indels 11; Gaps 3;

OY 27 DPNPCE--NGCICPLAVGSPGCPDGPDPNCS 61
Db 2 DPEKTCMSSGICIG-----DNCKC-----TTCNCKT 29

RESULT 12
O9ZH31 PRELIMINARY; PRT; 64 AA.
AC O9ZH31;
DT 01-MAY-1999 (TREMREL. 10, Created)
DR 01-MAY-1999 (TREMREL. 10, Last sequence update)
DR 01-MAY-2000 (TREMREL. 13, Last annotation update)
DS FXN-LIKE PROTEIN.
OS Acetabacter diazotrophicus.
OC Bacteriis; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconacetobacter.
CX NCBI_TaxID=33996;
NM [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PALS;
RA Teixeira K.R.S., Morgan T., Gallier R., Zellerman E.-M., Baldani J.I.,
RA Kennedy C., Meletzus D.;
RT "Molecular analysis of the chromosomal region encoding the nifA and
RT nifB genes of Acetobacter diazotrophicus."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF072689; AAC99366.1; -.
DR HSSP: P00208; 1BLU.
DR INTERPRO: IPR001450; -.
DR PFAM: PF00037; fer4; 1.
DR POSITE: PS00198; 4FEAS_FERREROXIN; 1.
RW Iron-sulfur.
SQ SEQUENCE 64 AA; 6852 MW; DDBE6E1C642F8ARC CRC64;

Query Match 13.7%; Score 49.5; DB 2; Length 64;
Best Local Similarity 36.8%; Pred. No. 58;
Matches 14; Conservative 2; Mismatches 19; Indels 3; Gaps 2;

OY 31 CENGICLPG--LAVGSFSC-ECPDGFTDPNCSVVEV 65
Db 20 CPNGAISLKTLDYVIAKKTCEGCFDTPCVCVCPV 57

RESULT 13
O73914 PRELIMINARY; PRT; 60 AA.
AC O73914;
DT 01-AUG-1998 (TREMREL. 07, Created)

```

01-AUG-1998 (TREMBLrel. 07, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE METALLOTHIONEIN (MT).
 GN MT.
 OS Notothenia coriiceps neglecta (Black rockcod) (Yellowbelly rockcod),
 OS and Chionodraco rastrospinosus.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
 CC Nototheniidae; Nototheniidae; Notothenia.
 OX NCBI_TaxID=8209, 34790;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Scudiero R., Verde C., Cargnale V., Capasso C., di Prisco G.,
 RL Parisi E.,
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RC TISSUE=LIVER;
 CC SEQUENCE FROM N.A.
 RA Cargnale V., Capasso C., Scudiero R., Parisi E.;
 RT "Metallothionein in Antarctic organisms."
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
 CC GLUCOCORTICOIDS.
 CC EMBL: AJ006484; CA07063.1; -;
 DR EMBL: AJ011584; CA09714.1; -;
 DR HSSP: P02802; IDFS.
 DR INTERPRO: IPR000006; -;
 DR INTERPRO: IPR003019; -;
 DR PRAM: PF00131; metalthio. 1.
 DR PRINTS: PR00860; MVTERTERATE.
 DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.
 DR KW Metal-binding; Metal-chiolate cluster; Chelation; Zinc.
 FT METAL 23 23 CLUSTER.
 FT METAL 26 26 CLUSTER.
 FT METAL 16 16 CLUSTER.
 FT METAL 18 18 CLUSTER.
 FT METAL 21 21 CLUSTER.
 FT METAL 30 30 CLUSTER.
 SQ SEQUENCE 60 AA; 6019 MW; E866E7155A2C424A CRC64;

Query Match 13.4%; Score 48.5; DB 13; Length 60;
 Best Local Similarity 30.6%; Pred. No. 71;
 Matches 15; Conservative 2; Mismatches 13; Indels 19; Gaps 3;

27 DPNPCENGICLPLGAVGSFSC-----CPDGFDPNCSS 61
 Db 2 DPCGSKSGTCNCG---GSCCTNCCKSKCKSCPCPSGCT--KCAS 45

RESULT 14
 ID 093609 PRELIMINARY; PRT; 60 AA.
 AC 093609;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE METALLOTHIONEIN (MT).
 GN MT.
 OS Trematomus bernacchii.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
 CC Nototheniidae; Nototheniidae; Trematomus.
 OX NCBI_TaxID=40690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cargnale V., Capasso C., Scudiero R., Parisi E.;
 RL TISSUE=LIVER;
 CC "Metallothionein in Antarctic organisms.";

Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
 CC GLUCOCORTICOIDS.
 CC EMBL: AJ011585; CA09715.1; -;
 DR HSSP: P02802; IDFS.
 DR INTERPRO: IPR000006; -;
 DR INTERPRO: IPR003019; -;
 DR PRAM: PF00131; metalthio. 1.
 DR PRINTS: PR00860; MVTERTERATE.
 DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.
 DR KW Metal-binding; Metal-chiolate cluster; Chelation; Zinc.
 FT METAL 23 23 CLUSTER.
 FT METAL 26 26 CLUSTER.
 FT METAL 16 16 CLUSTER.
 FT METAL 18 18 CLUSTER.
 FT METAL 21 21 CLUSTER.
 FT METAL 30 30 CLUSTER.
 SQ SEQUENCE 60 AA; 6018 MW; AC66F3015A2C4251 CRC64;

Query Match 13.4%; Score 48.5; DB 13; Length 60;
 Best Local Similarity 30.6%; Pred. No. 71;
 Matches 15; Conservative 2; Mismatches 13; Indels 19; Gaps 3;

27 DPNPCENGICLPLGAVGSFSC-----CPDGFDPNCSS 61
 Db 2 DPCGSKSGTCNCG---GSCCTNCCKSKCKSCPCPSGCT--KCAS 45

RESULT 15
 ID 013259 PRELIMINARY; PRT; 60 AA.
 AC 013259;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE METALLOTHIONEIN (MT).
 GN MTB.
 OS Chionodraco hamatus, Morone saxatilis (Striped bass), and
 OS Gymnodraco acuticeps.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
 CC Nototheniidae; Channichthyidae; Chionodraco.
 OX NCBI_TaxID=36188, 34816, 8218;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SPECIES=C.hamatus;
 RA Cargnale V., Scudiero R., Capasso A., Capasso C., Kille P.,
 RA di Prisco G., Parisi E.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA SPECIES=M.saxatilis;
 RA Leclerc G.M., Leclerc G.J., Ely B.;
 RT "Molecular cloning of the striped bass (Morone saxatilis)
 RT metallothionein gene (sbMT).";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA SPECIES=G.acuticeps, and M.saxatilis; TISSUE=LIVER;
 RA Cargnale V., Scudiero R., Capasso C., Parisi E.;
 RT "Metallothioneins in Antarctic fish."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
 CC TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
 CC GLUCOCORTICOIDS.
 CC EMBL: Y12811; CA073160.1; -;
 DR EMBL: AF091100; AAC62501.1; -;
 DR EMBL: AJ007561; CA07556.1; -;

DR HSSP; P02802; 1DFS.
DR INTERPRO; IPR000006; -.
DR INTERPRO; IPR003019; -.
DR PFAM; PF00131; metalthio; 1.
DR PRINTS; PR00860; MTVERTERATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Chelation; Zinc.
FT METAL 14 14 CLUSTER.
FT METAL 16 16 CLUSTER.
FT METAL 18 18 CLUSTER.
FT METAL 20 20 CLUSTER.
FT METAL 21 21 CLUSTER.
FT METAL 23 23 CLUSTER.
FT METAL 25 25 CLUSTER.
FT METAL 26 26 CLUSTER.
FT METAL 28 28 CLUSTER.
FT METAL 30 30 CLUSTER.
SQ SEQUENCE 60 AA; 5978 MW; 59665AC61BC424B CRC64;

Query Match 13.4%; Score 48.5; DB 13; Length 60;
Best Local Similarity 30.6%; Pred. No. 71;
Matches 15; Conservative 2; Mismatches 13; Indels 19; Gaps 3;
QY 27 DPNPCENGICLPGLAVGSFSCF-----CPDGFTDPNCS 61
DB 2 DPCDCSKSGTCNCG--GSCCTCNCSCCTSCCKSCPCPCPGCT--KCAS 45

Search completed: May 23, 2001, 08:34:40
Job time: 106 sec


```

PT Del-1 polypeptide sequences - useful for treatment of cancer,
XX diabetic retinopathy, rheumatoid arthritis and endometriosis
PS Claim 40; Fig 10; 73pp: English.
XX
CC The present sequence represents an epidermal growth factor like domain
CC from developmentally-regulated endothelial cell locus-1 (Del-1). Del-1
CC also has discolidin I/factor VIII-like domains. The Del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
SQ Sequence 57 AA:

Query Match 61.5%; Score 222.5; DB 20; Length 57;
Best Local Similarity 88.9%; Pred. No. 2.2e-13;
Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 26 CDNPENCGTCTGGLAVG-----SPSCPCPGPDPNCGSVVEV 65
Db 1 cdnpenggtclpglavgxxxxxsfscpcpgfdpncssvvev 45

RESULT 2
W0368
ID W0368 standard; Peptide; 25 AA.
XX
AC W0368;
XX
DT 03-MAY-1997 (first entry)
XX
DE Human Del-1 signal peptide.
XX
KW Del-1; developmentally-regulated endothelial cell locus-1;
KW signal transduction; cancer; tumour marker; angiogenesis;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW cardiac ischaemia; stroke; vascular disease; wound healing;
KW vlnnary; bone formation; diagnosis; therapy.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT Peptide 3-21
FT /label= Sig_peptide
XX
FM W09640769-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-0509456.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI: 1997-052233/05.
XX
PT New developmentally regulated endothelial cell locus-1 (del-1) gene
PT - used to develop prods for the diagnosis and treatment of cancer
PT and conditions involving abnormal angiogenesis
XX
PS Example; Fig 9; 137pp: English.

```

```

XX
CC The predicted amino acid sequence (W0368) at the N-terminus
CC of the human developmentally-regulated endothelial cell locus-1
CC (del-1) gene product (W0365) shows characteristics common to
CC signal peptides. The putative signal begins with a basic Arg
CC residue and is followed by a stretch of 18 hydrophobic residues.
CC The Chou and Fasman algorithm predicts that the putative signal
CC sequence is followed by a protein turn structure, a feature
CC commonly found after signal peptides. The Del-1 protein is
CC secreted by expressing cells.
XX
SQ Sequence 25 AA:

Query Match 35.1%; Score 127; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAVMLVGLSLGVPGFGKDI 25
Db 1 mkrsavawllvgslgvpfgkgdi 25

RESULT 3
W94698
ID W94698 standard; peptide; 25 AA.
XX
AC W94698;
XX
DT 05-MAY-1999 (first entry)
XX
DE Human Del-1 signal protein.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discolidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Homo sapiens.
XX
PN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI: 1999-189720/16.
XX
PT Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
PS Example; Column 63-65; 73pp: English.
XX
CC The present sequence is the human developmentally-regulated endothelial
CC cell locus 1 (Del-1) amino terminus peptide sequence. The Del-1 protein
CC has epidermal growth factor (EGF) like domains and discolidin I/factor
CC VIII-like domains. The Del-1 proteins have an inhibitory effect on
CC angiogenesis (blood vessel growth), this activity may be useful
CC clinically to prevent neovascularisation of tissues such as tumour
CC nodules and prevention of metastases. The anti-angiogenic activity of
CC Del-1 may be used to treat abnormal conditions that result from
CC angiogenesis, including cancer, diabetic retinopathy, rheumatoid
CC arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be
CC used to treat cardiac ischaemia, thrombotic stroke, wound healing and
CC peripheral vascular disease. Del-1 is also useful for promoting bone

```

CC formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis inducer.

XX Sequence 25 AA:

Query Match 35.1%; Score 127; DB 20; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAVMLVGLSLGVPFGKCDI 25
Db 1 mkrsavavmlvlgslgvpfgkcdi 25

RESULT 4

R96227 R96227 standard; Protein; 46 AA.

XX AC R96227;

XX DT 06-AUG-1996 (first entry)

XX DE Novel growth factor domain fragment D.

XX KW Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;

XX KM fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;

XX KW protein engineering; growth factor domain.

XX OS Synthetic.

XX PN US5504001-A.

XX PD 02-APR-1996.

XX FE 25-NOV-1987; 87US-0125629.

XX FR 25-NOV-1987; 87US-0125629.

XX FR 28-JAN-1992; 92US-0827587.

XX PR 06-JUN-1994; 94US-0254485.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Foster DC;

XX DR WPI; 1996-187699/19.

XX DR N-PSDB; T27592.

XX XX Hybrid plasminogen activator comprises human tPA activator and N-terminal crosslinking domain from alpha2-plasmin inhibitor -

XX PT useful to treat thrombosis and image blood clots

XX PS Example 6; Fig 15; 35pp; English.

XX CC Amino acid substitutions were designed in the growth factor

XX CC domain of tissue plasminogen activator (tPA) with the goal of

XX CC disrupting possible specific receptor interactions.

XX CC Oligonucleotides (T27589-92) encoding the growth factor region

XX CC replacement domains A-D (R96224-27) were generated. Fragment D

XX CC results in replacement of tPA amino acids 52-70 with a consensus

XX CC receptor binding region. Mutant tPAs were expressed in BHK cells

XX CC and characterised for plasma half life and fibrin binding properties.

XX SQ Sequence 46 AA:

Query Match 30.9%; Score 112; DB 17; Length 46;

Best Local Similarity 47.1%; Pred. No. 0.00121; Mismatches 11; Indels 0; Gaps 0;

QY 26 CDPNPGNGICLPGLAVGSGCECPDGTDPNC 59
Db 7 csepnpgngiclpplavgsfsccepdgtdpnc 40

RESULT 5

R96226 R96226 standard; Protein; 46 AA.

XX AC R96226;

XX DT 06-AUG-1996 (first entry)

XX DE Novel growth factor domain fragment C.

XX KW Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;

XX KM fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;

XX KW protein engineering; growth factor domain; Factor IX.

XX OS Synthetic.

XX PN US5504001-A.

XX PD 02-APR-1996.

XX FE 25-NOV-1987; 87US-0125629.

XX FR 25-NOV-1987; 87US-0125629.

XX FR 28-JAN-1992; 92US-0827587.

XX PR 06-JUN-1994; 94US-0254485.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Foster DC;

XX DR WPI; 1996-187699/19.

XX DR N-PSDB; T27591.

XX XX Hybrid plasminogen activator comprises human tPA activator and N-terminal crosslinking domain from alpha2-plasmin inhibitor -

XX PT useful to treat thrombosis and image blood clots

XX PS Example 6; Fig 15; 35pp; English.

XX CC Amino acid substitutions were designed in the growth factor

XX CC domain of tissue plasminogen activator (tPA) with the goal of

XX CC disrupting possible specific receptor interactions.

XX CC Oligonucleotides (T27589-92) encoding the growth factor region

XX CC replacement domains A-D (R96224-27) were generated. Fragment C

XX CC results in replacement of tPA amino acids 52-55 with the

XX CC corresponding 4 amino acids of human Factor IX. Mutant tPAs were

XX CC expressed in BHK cells and characterised for plasma half life and

XX CC fibrin binding properties.

XX SQ Sequence 46 AA:

Query Match 29.3%; Score 106; DB 17; Length 46;

Best Local Similarity 47.1%; Pred. No. 0.004; Mismatches 14; Indels 0; Gaps 0;

QY 26 CDPNPGNGICLPGLAVGSGCECPDGTDPNC 59
Db 7 csepnpgngiclpplavgsfsccepdgtdpnc 40

RESULT 6

W80576 W80576 standard; Protein; 52 AA.

XX AC W80576;

XX DT 16-DEC-1998 (first entry)

XX DE Human factor VII EGF domain primary sequence.

KW O-fucosyltransferase; epidermal growth factor; EGF; glycosylation;
 KM O-fucose; inhibitor; sensory neuron; retinal neuron; human; heart;
 XX
 OS Synthetic.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 44..49
 FT /note= "histidine tag"
 XX
 PN MO9833924-A1.
 PD 06-AUG-1998.
 PF 17-DEC-1997; 97MO-US23401.
 XX
 PR 26-NOV-1997; 97US-0978741.
 PR 31-JAN-1997; 97US-0792498.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Spellman MW, Wang Y;
 XX
 DR WPI: 1998-437477/37.
 XX
 PT Human O-fucosyltransferase able to glycosylate epidermal growth
 PT factor domains - useful for diagnosis and treatment of diseases
 PT Involving overexpression of the enzyme
 XX
 PS Example 2; Page 39; 90pp; English.
 XX
 CC This represents the primary sequence of the first EGF domain from human
 CC factor VII contained in a plasmid construct. The invention provides a
 CC human heart O-fucosyltransferase enzyme that can glycosylate an epidermal
 CC growth factor (EGF) domain of a polypeptide with an activated O-fucose
 CC residue. Inhibitors of O-fucosyltransferase, e.g. mutants with increased
 CC affinity for the EGF domains, are used in diagnosis and treatment of
 CC conditions associated with overexpression of O-fucosyltransferase to
 CC promote survival of sensory (retinal) neurons. Probes based on EGF domain
 CC polypeptide are used to detect gene amplification and expression. The
 CC expression can also be determined at the protein level using antibodies
 CC specific for O-fucosyltransferase.
 CC
 SO Sequence 52 AA:
 XX
 XX
 Query Match 27.9%; Score 101; DB 19; Length 52;
 Best Local Similarity 48.6%; Pred. No. 0.012; Mismatches 2; Gaps 1;
 Matches 18; Conservative 3; Indels 14;
 OY 23 GDIDPNCENGICLPGLAVGFSCECPDGFDPNC 59
 DB 4 gdcgeapclngskcd--lmsyewcpfigeknc 38
 XX
 XX
 RESULT 7
 W94688 W94688 standard; peptide; 42 AA.
 AC W94688:
 XX
 XX
 DT 05-MAY-1999 (first entry)
 XX
 DE Del-1 epidermal growth factor like domain #3.
 XX
 KM Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KM discoidin I; factor VII-like domain; epidermal growth factor; EGF;
 KM diabetic retinopathy; rheumatoid arthritis; endometriosis;
 XX
 OS Mus sp.
 OS Synthetic.

FN US5877281-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 05-JUN-1996; 96US-0659235.
 XX
 PR 05-JUN-1996; 96US-0659235.
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Hogan B, Quarterman T, Snodgrass HR, Zupancic TJ;
 XX
 DR WPI: 1999-189720/16.
 XX
 PT Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Claim 40; Fig 10; 73pp; English.
 XX
 CC The present sequence represents an epidermal growth factor like domain
 CC from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
 CC also has discoidin I/factor VII-like domains. The Del-1 proteins
 CC have an inhibitory effect on angiogenesis (blood vessel growth), this
 CC activity may be useful clinically to prevent neovascularisation of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC Integrin, and is an apoptosis inducer.
 CC
 SO Sequence 42 AA:
 XX
 XX
 Query Match 25.8%; Score 93.5; DB 20; Length 42;
 Best Local Similarity 43.6%; Pred. No. 0.048; Mismatches 5; Gaps 1;
 Matches 17; Conservative 5; Indels 12;
 OY 26 CDPNCEGICLPGLA-----VGSFCECPDGFDPNC 59
 DB 1 cevepekngicldlvaxxxxxxxxxxscpcpelmgrnc 39
 XX
 XX
 RESULT 8
 R96224 R96224 standard; Protein; 46 AA.
 AC R96224:
 XX
 DT 06-AUG-1996 (first entry)
 XX
 DE Novel growth factor domain fragment A.
 XX
 KM Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
 KM fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
 KM protein engineering; growth factor domain.
 XX
 OS Synthetic.
 XX
 PN US5504001-A.
 PD 02-APR-1996.
 XX
 PF 25-NOV-1987; 87US-0125629.
 XX
 PR 25-NOV-1987; 87US-0125629.
 PR 28-JAN-1992; 92US-0827587.
 PR 06-JUN-1994; 94US-0254485.
 XX

PA (ZYMO) ZYMOGENETICS INC.
 PI Foster DC;
 DR WPI: 1996-187699/19.
 DR N-PSDB: T27589.
 XX
 XX Hybrid plasminogen activator comprises human tPA activator and
 PT N-terminal crosslinking domain from alpha2-plasmin inhibitor
 PT useful to treat thrombosis and image blood clots
 XX
 PS Example 6; Fig 15; 35pp; English.
 CC Amino acid substitutions were designed in the growth factor
 CC domain of tissue plasminogen activator (tPA) with the goal of
 CC disrupting possible specific receptor interactions.
 CC Oligonucleotides (T27589-92) encoding the growth factor region
 CC replacement domains A-D (R96224-27) were generated. Fragment A
 CC results in replacement of tPA amino acids 63-70 with a consensus
 CC receptor binding region. Mutant tPAs were expressed in BHK cells
 CC and characterised for plasma half life and fibrin binding properties.
 CC
 XX Sequence 46 AA:
 SO
 Query Match 25.7%; Score 93; DB 17; Length 46;
 Best Local Similarity 41.2%; Pred. No. 0.057;
 Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;
 QY 26 CDPNPGNGGICLPGLAVGSFSCPCPGFTDPNC 59
 DB 7 cseprcfnqgtcmegnhlanfvcqcgqefgksc 40
 ID B32422 standard; Protein: 44 AA.
 AC R96225:
 XX R96225:
 AC
 XX 06-AUG-1996 (first entry)
 DT
 XX
 DE Novel growth factor domain fragment B.
 XX
 XX Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;
 KM fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;
 KM protein engineering; growth factor domain; Factor IX.
 XX
 OS Synthetic.
 XX
 XX US5504001-A.
 PN
 XX 02-APR-1996.
 PD
 XX 25-NOV-1987; 87US-0125629.
 PF
 XX 25-NOV-1987; 87US-0125629.
 XX 28-JAN-1992; 92US-0827587.
 PR 06-JUN-1994; 94US-0254485.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Foster DC;
 PI
 XX WPI: 1996-187699/19.
 DR N-PSDB: T27590.
 DR
 XX Hybrid plasminogen activator comprises human tPA activator and
 PT N-terminal crosslinking domain from alpha2-plasmin inhibitor
 PT useful to treat thrombosis and image blood clots
 XX
 PS Example 6; Fig 15; 35pp; English.

CC Amino acid substitutions were designed in the growth factor
 CC domain of tissue plasminogen activator (tPA) with the goal of
 CC disrupting possible specific receptor interactions.
 CC Oligonucleotides (T27589-92) encoding the growth factor region
 CC replacement domains A-D (R96224-27) were generated. Fragment B
 CC results in replacement of tPA amino acids 52-91 with the entire
 CC growth factor region of human factor IX. Mutant tPAs were
 CC expressed in BHK cells and characterised for plasma half life and
 CC fibrin binding properties.
 CC
 XX Sequence 44 AA:
 SO
 Query Match 25.4%; Score 92; DB 17; Length 44;
 Best Local Similarity 47.1%; Pred. No. 0.067;
 Matches 16; Conservative 3; Mismatches 13; Indels 2; Gaps 1;
 QY 26 CDPNPGNGGICLPGLAVGSFSCPCPGFTDPNC 59
 DB 7 csepnclngysckd--lasyecwcpigfegknc 38
 ID B32422 standard; Protein: 44 AA.
 AC B32422:
 XX
 XX 16-JAN-2001 (first entry)
 DT
 XX
 DE Human secreted protein sequence encoded by gene 2 SEQ ID NO:108.
 XX
 XX Human: secreted protein; cytostatic; immunostimulant; antiproliferative;
 KW cardiant; antiarrhythmic; antiviral; antibacterial; antifungal; cancer;
 KW antiparasitic; neuroprotective; nootropic; antiinflammatory; anti-HIV;
 KW antiangiogenic; antiarteriosclerotic; diagnosis; immune disorder; AIDS;
 KW autoimmune disease; haematopoietic cell disorder; blood protein disorder;
 KW agammaglobulinemia; hyperproliferative disease; Gaucher's disease;
 KW cardiovascular disorder; congenital heart defect; pulmonary atresia;
 KW arhythmia; ischaemia; angiogenesis related disorder; Crohn's disease;
 KW atherosclerosis; neurological disease; Alzheimer's disease;
 KW Huntington's; infectious disease; cat-scratch disease.
 XX
 XX Homo sapiens.
 OS
 XX WO200047602-A1.
 PN
 XX 17-AUG-2000.
 PD
 XX 08-FEB-2000; 2000WO-US03062.
 PF
 XX 10-FEB-1999; 99US-0119468.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM, Ehner R, Young PE, Ni J, Soppet DR, Moore PA;
 PI Shi Y, Lafleur DW, Olsen HS, Florence KA, Komatsu G;
 DR WPI: 2000-543578/49.
 XX
 XX New human nucleic acids encoding secreted proteins, useful in the
 PT treatment, prevention or diagnosis of immune disorders (e.g. autoimmune
 PT diseases), blood protein disorders and hyperproliferative diseases
 PT (e.g. Gaucher's disease)
 XX
 XX Disclosure; Page 9; 48pp; English.
 CC The polynucleotide sequences given in C55190 to C55235 encode the human
 CC secreted proteins given in B32371 to B32484. Human secreted proteins
 CC have activities based on the tissues and cells the genes are expressed
 CC in. Examples of activities include: cytostatic; immunostimulant;
 CC antiproliferative; cardiant; antiarrhythmic; antiviral; antibacterial;
 CC antifungal; antiparasitic; neuroprotective; nootropic; antiinflammatory;

CC antiangiogenic; anti-HIV; and antiarteriosclerotic. The polynucleotides
 CC and polypeptides, or their agonists and antagonists, can be used for
 CC treating, preventing or diagnosing immune disorders (e.g. cancer,
 CC autoimmune diseases), disorders of haematopoietic cells, blood protein
 CC disorders (e.g. agammaglobulinemia), hyperproliferative diseases
 CC (e.g. Gaucher's disease), cardiovascular disorders (e.g. congenital
 CC heart defects, pulmonary atresia, arrhythmias, ischemia), angiogenesis
 CC related disorders (e.g. Crohn's disease, atherosclerosis), neurological
 CC diseases (e.g. Alzheimer's disease, Huntington's chorea), infectious
 CC diseases (e.g. AIDS, cat-scratch disease and other bacterial, viral,
 CC parasitic or fungal diseases) C5181 to C5189 and B32370 represent
 CC sequences used in the exemplification of the present invention.

XX Sequence 44 AA;

Query Match 24.38; Score 88; DB 21; Length 44;
 Best Local Similarity 50.0%; Pred. No. 0.15;
 Matches 17; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

OY 26 CDPNCPENGICLPGLAVGSFCECPDGFDPNC 59
 || ||| ||| | || | || |
 Db 4 cdgnpcangsc--setprfctcprgyjrc 35

RESULT 11

W94687 standard; peptide; 43 AA.

XX W94687;

DT 05-MAY-1999 (first entry)

XX Del-1 epidermal growth factor like domain #2.

XX Del-1: developmentally-regulated endothelial cell locus 1; cancer;
 KW discoidin II factor VIII-like domain; epidermal growth factor; EGF;
 KM diabetic retinopathy; rheumatoid arthritis; endometriosis;
 XX angiogenesis.

OS Mus sp.
 OS Synthetic.

PN US5877281-A.

PD 02-MAR-1999.

PF 05-JUN-1996; 96US-0659235.

PR 05-JUN-1996; 96US-0659235.

PR 07-JUN-1995; 95US-0480229.

PA (PROG-) PROGENITOR INC.

PA (UYVA-) UNIV VANDERBILT.

PI Hogan B, Quertemous T, Snodgrass HR, Zupancic TJ;

DR WPI; 1999-189720/16.

XX Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis

XX Claim 40; Fig 10; 73pp; English.

XX The present sequence represents an epidermal growth factor like domain
 CC from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
 CC also has discoidin I/factor VIII-like domains. The Del-1 proteins
 CC have an inhibitory effect on angiogenesis (blood vessel growth), this
 CC activity may be useful clinically to prevent neovascularisation of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes

CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC integrin, and is an apoptosis inducer.

XX Sequence 43 AA;

Query Match 24.29; Score 87.5; DB 20; Length 43;
 Best Local Similarity 41.0%; Pred. No. 0.17;
 Matches 16; Conservative 3; Mismatches 15; Indels 5; Gaps 1;

OY 26 CDPNCPENGIC-----LGLAVGSFCECPDGFDPNC 59
 | |||| ||| | : || | : |
 Db 1 cdnphngtcelseayrdfllgyvckcprgfhnc 39

RESULT 12

Y27028 standard; Protein; 38 AA.

XX Y27028;

DT 08-OCT-1999 (first entry)

DE Amino acid sequence of rat cubilin EGF8 repeat.

XX Cubilin: epithelial glycoprotein receptor; vitamin B12; endocytosis;

KW toxicity; kidney; spleen; brain; liver; heart; thyroid; abortion; rat;

KW proteinuria; fetal malformation; fetal development; kidney damage; EGF.

OS Rattus sp.

PN W0933757-A1.

XX 29-JUL-1999.

PF 21-JAN-1999; 99WO-US01259.

PR 22-JAN-1998; 98US-0072197.

PA (INRM) INST NAT SANTE & RECH MEDICALE.

PA (TULA) TULANE EDUCATIONAL FUND.

PI Hammond TG, Verroust PJ;

DR WPI; 1999-479045/40.

XX New DNA encoding cubilin, used for treating toxicity, particularly

PT nephrotoxicity, and as marker of kidney damage

XX Example 18; Fig 6A; 135pp; English.

XX The invention relates to a rat cubilin protein. Cubilin is a ligand-
 CC binding, epithelial glycoprotein receptor that facilitates uptake of
 CC intrinsic factor/vitamin B12 complexes in intestines and kidney. It is
 CC also involved in endocytosis and trafficking of light immunoglobulin
 CC chains in renal proximal tubule cells. Host cells containing a vector
 CC comprising the rat cubilin DNA sequence can be used for the recombinant
 CC expression of the protein. Cubilin and its fragments, are used to treat
 CC or reduce toxicity, particularly in kidneys, spleen, brain, liver, heart
 CC and thyroid. Cubilin mutations may also be implicated in idiopathic
 CC abortions, fetal malformation, poor fetal development and spontaneous
 CC proteinuria. Cubilin may also be used to raise specific antibodies, used
 CC for its detection, or clones that express it, in standard immunoassays.
 CC Fragments of cubilin DNA can also be used to detect cubilin mRNA in cell
 CC and tissues, by hybridization. Abnormal levels of cubilin in the urine
 CC are indicative of kidney damage. Sequences Y27021-35 represent EGF
 CC repeat sequences of rat cubilin and from homologous regions of other
 CC proteins.

XX Sequence 38 AA;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:32:28 ; Search time 36.45 Seconds
(without alignments)
286.846 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Sequence: 1 MKRSYAWMLVGLSLGVPF.....FSCBCPDGTFDPMCSYVEV 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 480680

Minimum DB seq length: 0

Maximum DB seq length: 65

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	222.5	61.5	57	US-08-480-2298-23	Sequence 23, Appl
2	222.5	61.5	57	US-08-659-2358-23	Sequence 23, Appl
3	222.5	61.5	57	US-09-237-981-23	Sequence 23, Appl
4	127	35.1	25	US-08-480-2298-22	Sequence 22, Appl
5	127	35.1	25	US-08-659-2358-22	Sequence 22, Appl
6	127	35.1	25	US-09-237-981-22	Sequence 22, Appl
7	113.5	31.4	44	US-60-196-110-6855	Sequence 6855, Ap
8	110.5	30.5	50	US-60-192-739-2813	Sequence 2813, Ap
9	110.5	30.5	50	US-60-194-243-2012	Sequence 2012, Ap
10	110	30.4	33	US-60-140-956-2007	Sequence 2007, Ap

11	110	30.4	55	US-60-162-243-724	Sequence 724, App
12	110	30.4	55	US-60-162-247-5176	Sequence 5176, App
13	110	30.4	57	US-60-170-373-2725	Sequence 2725, Ap
14	109	30.1	39	US-60-192-739-3764	Sequence 3764, Ap
15	109	30.1	42	US-08-968-800-5	Sequence 5, Appl1
16	109	30.1	42	US-09-157-308-7	Sequence 7, Appl1
17	109	30.1	42	US-09-249-697-7	Sequence 7, Appl1
18	109	30.1	42	US-09-363-316B-7	Sequence 7, Appl1
19	109	30.1	42	US-09-400-010-7	Sequence 7, Appl1
20	109	30.1	42	US-09-687-860-7	Sequence 7, Appl1
21	109	30.1	42	US-09-192-739-2815	Sequence 2815, Ap
22	105	29.0	61	US-60-194-243-2014	Sequence 2014, Ap
23	105	29.0	32	US-60-160-203-3896	Sequence 3896, Ap
24	103	28.5	32	US-60-160-203-3897	Sequence 3897, Ap
25	102	28.2	42	US-60-169-840-5748	Sequence 5748, Ap
26	102	28.2	60	US-60-162-247-5136	Sequence 5136, Ap
27	102	28.2	60	US-60-236-356-16205	Sequence 16205, Ap
28	101	27.9	38	US-09-328-925-8	Sequence 8, Appl1
29	101	27.9	38	US-08-332-498A-15	Sequence 15, Appl1
30	101	27.9	32	US-09-333-728A-15	Sequence 15, Appl1
31	101	27.9	52	US-09-333-728A-15	Sequence 15, Appl1
32	101	27.9	52	US-09-333-728A-15	Sequence 15, Appl1
33	101	27.9	52	US-09-333-728A-15	Sequence 15, Appl1
34	101	27.9	52	US-09-333-728A-15	Sequence 15, Appl1
35	99.5	27.5	38	US-60-169-840-4880	Sequence 4880, Ap
36	99.5	27.5	38	US-60-169-840-4880	Sequence 4880, Ap
37	99	27.3	34	US-60-164-762-7750	Sequence 7750, App
38	98	27.1	40	US-60-236-359-18942	Sequence 18942, A
39	97	26.8	57	US-60-162-247-3304	Sequence 3304, Ap
40	97	26.8	57	US-60-162-247-3304	Sequence 3304, Ap
41	96	26.5	32	US-60-198-954-775	Sequence 775, App
42	96	26.5	32	US-60-160-203-3264	Sequence 3264, Ap
43	96	26.5	32	US-60-163-123-1461	Sequence 1461, Ap
44	96	26.5	40	US-60-186-656-802	Sequence 802, App
45	96	26.5	40	US-60-190-000-146	Sequence 146, App
			23	US-60-160-203-5445	Sequence 5445, Ap

ALIGNMENTS

RESULT 1

US-08-480-2298-23

Sequence 23, Application US/084802298

GENERAL INFORMATION:

APPLICANT: Quentemus, Thomas

APPLICANT: Hogan, Bridgid

APPLICANT: Snodgrass, R. Ralph

APPLICANT: Zupancic, Thomas J

TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

TITLE OF INVENTION: CELL LOCUS-1

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESS: Pennine & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/480,2298

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Polissant, Brian M.

REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-026

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229B-23

Query Match 61.5%: Score 222.5; DB 8; Length 57;
Best Local Similarity 88.9%: Pred. No. 7.9e-17;
Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 26 CDPNCEGICLPLGLAVG-----SFCSCPCDGTDPNCSVVEY 65
DB 1 CDPNCEGICLPLGLAVGXXXXXSFSCPCDGTDPNCSVVEY 45

RESULT 2
US-08-659-235B-23
Sequence 23, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF INVENTIONS: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235B-23

Query Match 61.5%: Score 222.5; DB 10; Length 57;
Best Local Similarity 88.9%: Pred. No. 7.9e-17;
Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 26 CDPNCEGICLPLGLAVG-----SFCSCPCDGTDPNCSVVEY 65
DB 1 CDPNCEGICLPLGLAVGXXXXXSFSCPCDGTDPNCSVVEY 45

RESULT 3
US-09-237-981-23
Sequence 23, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF INVENTIONS: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-237-981-23

Query Match 61.5%: Score 222.5; DB 16; Length 57;
Best Local Similarity 88.9%: Pred. No. 7.9e-17;
Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 26 CDPNCEGICLPLGLAVG-----SFCSCPCDGTDPNCSVVEY 65
DB 1 CDPNCEGICLPLGLAVGXXXXXSFSCPCDGTDPNCSVVEY 45

RESULT 4
US-08-480-229B-22
Sequence 22, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF INVENTIONS: CELL LOCUS-1
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,223B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229B-22

Query Match 35.1%; Score 127; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 7,1e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVMLVGLSLGVPQGRGDI 25
DB 1 MKRSYAVMLVGLSLGVPQGRGDI 25

RESULT 5
US-08-659-235B-22
Sequence 22, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quertemous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235B-22

Query Match 35.1%; Score 127; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 7,1e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVMLVGLSLGVPQGRGDI 25
DB 1 MKRSYAVMLVGLSLGVPQGRGDI 25

RESULT 6
US-09-237-981-22
Sequence 22, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quertemous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-237-981-22

Query Match 35.1%; Score 127; DB 16; Length 25;
Best Local Similarity 100.0%; Pred. No. 7,1e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKRSVAWLVGLVSLGVPQFGKDI 25
DB 1 MKRSVAWLVGLVSLGVPQFGKDI 25

RESULT 7
US-60-196-710-6855
Sequence 6855, Application US/60196710
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
AND USES THEREOF
FILE REFERENCE: CL000450
CURRENT APPLICATION NUMBER: US/60/196,710
CURRENT FILING DATE: 2000-04-13
NUMBER OF SEQ ID NOS: 7166
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6855
LENGTH: 44
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(44)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-196-710-6855

Query Match
Best Local Similarity 73.3%; Pred. No. 3.6e-05;
Matches 22; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 36 ICPLGLAVGSCPCPDGFTDPNCSSYEV 65
DB 15 VCDWLMV-PPCCPCPDGFTDPNCSSYEV 43

RESULT 8
US-60-192-739-2813
Sequence 2813, Application US/60192739
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
USES THEREOF
FILE REFERENCE: CL000406
CURRENT APPLICATION NUMBER: US/60/192,739
CURRENT FILING DATE: 2000-03-28
NUMBER OF SEQ ID NOS: 4532
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2813
LENGTH: 50
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(55)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-192-739-2813

Query Match
Best Local Similarity 52.6%; Pred. No. 8.6e-05;
Matches 20; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

QY 26 CDPNCPENGICLPLGLAVGSCPCPDGFTDPNCSSYV 63
DB 1 CASHPCONGSCCEPGLHSG-FTCSCPDShTGRCEMVV 37

RESULT 9
US-60-194-243-2012
Sequence 2012, Application US/60194243
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
USES THEREOF
FILE REFERENCE: CL000417
CURRENT APPLICATION NUMBER: US/60/194,243
CURRENT FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 3242
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2012
LENGTH: 50
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(55)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-194-243-2012

Query Match
Best Local Similarity 52.6%; Pred. No. 8.6e-05;
Matches 20; Conservative 3; Mismatches 14; Indels 1; Gaps 1;

QY 26 CDPNCPENGICLPLGLAVGSCPCPDGFTDPNCSSYV 63
DB 1 CASHPCONGSCCEPGLHSG-FTCSCPDShTGRCEMVV 37

RESULT 10
US-60-140-956-2007
Sequence 2007, Application US/60140956
GENERAL INFORMATION:
APPLICANT: Keriavage, Anthony
TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
FILE REFERENCE: CL000041
CURRENT APPLICATION NUMBER: US/60/140,956
CURRENT FILING DATE: 1999-06-25
NUMBER OF SEQ ID NOS: 2638
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2007
LENGTH: 33
TYPE: PRT
ORGANISM: Drosophila
US-60-140-956-2007

Query Match
Best Local Similarity 63.3%; Pred. No. 6.4e-05;
Matches 19; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

QY 26 CDPNCPENGICLPLGLAVGSCPCPDGFT 55
DB 5 CSPNCPNRNGICLDG--DGFTECGMSGWT 32

RESULT 11
US-60-162-243-724
Sequence 724, Application US/60162243
GENERAL INFORMATION:
APPLICANT: Bonazzi, Vivien
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
USES THEREOF
FILE REFERENCE: CL000132
CURRENT APPLICATION NUMBER: US/60/162,243
CURRENT FILING DATE: 1999-10-29
NUMBER OF SEQ ID NOS: 828
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 724
LENGTH: 55
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(55)
OTHER INFORMATION: Xaa = Any Amino Acid
US-60-162-243-724
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:32:29 ; Search time 6.6 Seconds
(without alignments)
70.595 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97
Perfect score: 362
Sequence: 1 MKRSVAVMLVGLSGVPOF.....FSCECPDGFDPNCSVVEV 65

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41833 seqs, 7168127 residues

Total number of hits satisfying chosen parameters: 18580

Minimum DB seq length: 0
Maximum DB seq length: 65

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA New:*
1: /cgn2_6/ProdData/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ProdData/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ProdData/1/paa/US08_NEW_COMB.pep:*
4: /cgn2_6/ProdData/1/paa/US09_NEW_COMB.pep:*
5: /cgn2_6/ProdData/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ProdData/1/paa/US06_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	30.1	42	US-09-249-697A-7	Sequence 7, Appli
2	92	25.4	40	US-09-467-997-2	Sequence 2, Appli
3	88	24.3	48	US-09-802-704-11	Sequence 11, Appli
4	85	23.3	43	US-09-467-997-5	Sequence 5, Appli
5	82.3	22.8	39	US-09-467-997-3	Sequence 4, Appli
6	81	22.4	41	US-09-802-704-12	Sequence 3, Appli
7	60	16.6	21	US-09-802-704-11	Sequence 51, Appli
8	53.5	14.8	41	US-09-750-972-51	Sequence 5188, Ap
9	48.5	13.4	60	PCT-US01-01339-5188	Sequence 17, Appli
10	47	13.0	36	US-09-802-704-11	Sequence 450, Appl
11	46.5	12.7	58	US-09-817-427-450	Sequence 11, Appli
12	44.5	12.3	40	US-09-249-697A-11	Sequence 53, Appli
13	44.5	12.3	43	US-09-750-972-53	Sequence 22, Appli
14	44.5	12.3	43	US-09-802-704-14	Sequence 11, Appli
15	44	12.2	11	US-09-802-704-8	Sequence 14, Appli
16	44	12.2	19	US-09-802-704-13	Sequence 13, Appli
17	44	12.2	16	US-09-802-704-16	Sequence 16, Appli
18	44	12.2	27	US-09-817-427-613	Sequence 613, Appl
19	44	12.2	50	US-09-802-704-13	Sequence 1026, Ap
20	43.5	12.0	45	PCT-US01-11988-1026	Sequence 1026, Ap
21	43.5	12.0	45	PCT-US01-11988-1026	Sequence 1026, Ap
22	43.5	12.0	45	US-09-833-245-1026	Sequence 1028, Ap
23	43.5	12.0	45	US-09-833-245-1026	Sequence 1028, Ap
24	43	11.9	42	US-09-804-156-27	Sequence 27, Appli
25	42.5	11.7	48	US-09-801-436-316	Sequence 316, Appl
26	42	11.6	43	US-09-633-300-6	Sequence 6, Appli
27	42	11.6	49	PCT-US01-11501-253	Sequence 253, Appl

28	41.5	11.5	64	1	PCT-US01-01339-4801	Sequence 4801, Ap
29	41.5	11.5	65	1	PCT-US01-11988-660	Sequence 660, App
30	41.5	11.5	65	5	US-09-833-245-660	Sequence 660, App
31	41	11.3	39	5	US-09-249-697A-8	Sequence 8, Appli
32	41	11.3	45	5	US-09-249-697A-10	Sequence 10, Appli
33	40.5	11.2	52	5	US-09-836-377-340	Sequence 340, Appl
34	40	11.0	37	1	PCT-US01-11988-1027	Sequence 1027, Ap
35	40	11.0	37	5	US-09-833-245-1027	Sequence 1027, Ap
36	40	11.0	37	6	US-60-278-983-34	Sequence 34, Appli
37	40	11.0	39	5	PCT-US00-01979A-400	Sequence 400, Appl
38	40	11.0	39	5	US-09-493-795A-400	Sequence 400, Appl
39	40	11.0	40	5	US-09-804-156-26	Sequence 26, Appli
40	40	11.0	42	5	US-09-750-972-48	Sequence 48, Appli
41	40	11.0	44	1	PCT-US01-01351-233	Sequence 233, Appl
42	39	10.8	34	5	US-09-722-329-148	Sequence 148, App
43	39	10.8	44	5	US-09-801-436-177	Sequence 177, App
44	39	10.8	46	1	PCT-US01-01339-3492	Sequence 3492, Ap
45	39	10.8	51	1	PCT-US01-01339-4964	Sequence 4964, Ap

ALIGNMENTS

```
RESULT 1
US-09-249-697A-7
; Sequence 7, Application US/09249697A
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Yeung, George
; TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL
; FILE REFERENCE: 24011-727
; CURRENT APPLICATION NUMBER: US/09/249,697A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Drosophila Melanogaster
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(42)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-249-697A-7

Query Match 30.1%; Score 109; DB 5; Length 42;
Best Local Similarity 36.2%; Pred. No. 1e-06; 12; Indels 0; Gaps 0;
Matches 18; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 24 DTDPPNCEGICIPGLAVGSFSCCECPDGF 55
Db 2 DECKSNPCOMGSCXXDXVGSYXCXCPDGF 33

RESULT 2
US-09-467-997-2
; Sequence 2, Application US/09467997
; GENERAL INFORMATION:
; APPLICANT: Kitalowski, Jan
; APPLICANT: Uytendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 40
; TYPE: PRT
```

Wed May 23 08:50:33 2001

us-09-237-981-14_copy_33_97.clo.rapn

Page 2

ORGANISM: mouse
US-09-467-997-2

Query Match 25.4%; Score 92; DB 5; Length 40;
Best Local Similarity 50.0%; Pred. No. 8.9e-05;
Matches 19; Conservative 3; Mismatches 12; Indels 4; Gaps 2;

Db 4 DCDLGNRCHEAGKCLNTL--GSFECQCLQGTGPGC 39

RESULT 3
US-09-802-704-11
Sequence 11, Application US/09802704
GENERAL INFORMATION:
APPLICANT: Godbole, Shubhada
APPLICANT: Boyle, Bryan J
APPLICANT: Arterburn, Matthew C
APPLICANT: Liu, Chenghua
APPLICANT: Tang, Y Tom
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AND
FILE REFERENCE: HYS-40
CURRENT APPLICATION NUMBER: US/09/802,704
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/496,914
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 11
LENGTH: 48
TYPE: PRT
ORGANISM: Homo sapiens
US-09-802-704-11

Query Match 24.3%; Score 88; DB 5; Length 48;
Best Local Similarity 44.1%; Pred. No. 0.00011;
Matches 15; Conservative 3; Mismatches 14; Indels 2; Gaps 1;

Db 16 CSSQPCQNGTCEG--VNOIRICICPGRTGNRC 47

RESULT 4
US-09-467-997-5
Sequence 5, Application US/09467997
GENERAL INFORMATION:
APPLICANT: Klatjewski, Jan
APPLICANT: Uytendaele, Hendrik
TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
FILE REFERENCE: 53863-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/467,997
CURRENT FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 5
LENGTH: 43
TYPE: PRT
ORGANISM: mouse
US-09-467-997-5

Query Match 23.5%; Score 85; DB 5; Length 43;
Best Local Similarity 46.9%; Pred. No. 0.00061;
Matches 15; Conservative 5; Mismatches 10; Indels 2; Gaps 1;

QY 28 PNPENGICLPLGLAVGFSCECPDGTDPNC 59
DB 13 PSPCEHGSGCIN--TPGSFNCILCPYTGSRNC 42

RESULT 5
US-09-467-997-4
Sequence 4, Application US/09467997
GENERAL INFORMATION:
APPLICANT: Klatjewski, Jan
APPLICANT: Uytendaele, Hendrik
TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
FILE REFERENCE: 53863-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/467,997
CURRENT FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 4
LENGTH: 39
TYPE: PRT
ORGANISM: mouse
US-09-467-997-4

Query Match 22.8%; Score 82.5; DB 5; Length 39;
Best Local Similarity 48.4%; Pred. No. 0.0011;
Matches 15; Conservative 4; Mismatches 9; Indels 3; Gaps 1;

Db 11 NPCEHLGR--VTGGSFLDCCGRGTGPRC 38

RESULT 6
US-09-467-997-3
Sequence 3, Application US/09467997
GENERAL INFORMATION:
APPLICANT: Klatjewski, Jan
APPLICANT: Uytendaele, Hendrik
TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
FILE REFERENCE: 53863-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/467,997
CURRENT FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 3
LENGTH: 41
TYPE: PRT
ORGANISM: mouse
US-09-467-997-3

Query Match 22.4%; Score 81; DB 5; Length 41;
Best Local Similarity 40.0%; Pred. No. 0.0017;
Matches 14; Conservative 6; Mismatches 13; Indels 2; Gaps 1;

Db 8 MANSNCEHAGKCVN--TDGAFHCECLKGYAGPRC 40

RESULT 7
US-09-802-704-12
Sequence 12, Application US/09802704
GENERAL INFORMATION:
APPLICANT: Godbole, Shubhada
APPLICANT: Boyle, Bryan J
APPLICANT: Arterburn, Matthew C
APPLICANT: Liu, Chenghua
APPLICANT: Tang, Y Tom
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AND
FILE REFERENCE: HYS-40
CURRENT APPLICATION NUMBER: US/09/802,704
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/496,914
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 11
LENGTH: 48
TYPE: PRT
ORGANISM: Homo sapiens
US-09-802-704-12

FILE REFERENCE: HVS-40
CURRENT APPLICATION NUMBER: US/09/802,704
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 21
TYPE: PRT
ORGANISM: Homo sapiens
US-09-802-704-12

Query Match 16.6%; Score 60; DB 5; Length 21;
Best Local Similarity 47.6%; Pred. No. 0.23;
Matches 10; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY 30 PCNGGICLPGIAGVSGSCPC 50
DB 3 PCNGGTCVGG--VNGYRCIC 21

RESULT 8
US-09-750-972-51
Sequence 51, Application US/09750972
GENERAL INFORMATION:
APPLICANT: Pramod K. Srivastava
TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK
TITLE OF INVENTION: PROTEIN RECEPTOR AND USES THEREOF
FILE REFERENCE: 8449-134
CURRENT APPLICATION NUMBER: US/09/750,972
CURRENT FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: 09/750,972
PRIOR FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: 09/666,724
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 57
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 51
LENGTH: 41
TYPE: PRT
ORGANISM: Homo sapiens
US-09-750-972-51

Query Match 14.8%; Score 53.5; DB 5; Length 41;
Best Local Similarity 26.8%; Pred. No. 2.5;
Matches 11; Conservative 8; Mismatches 19; Indels 3; Gaps 1;

QY 25 ICDPN--PCNGGICLPGIAGVSGSCPCPDGFTDPNCSY 62
DB 1 VCDPSVFGCKDSARCISRAWCDGDNDCEINDSENCESI 41

RESULT 9
PCT-US01-01339-5188
Sequence 5188, Application PC/TUS0101339
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006PCT
CURRENT APPLICATION NUMBER: PCT/US01/01339
CURRENT FILING DATE: 2001-03-17
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5188
LENGTH: 60
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-01339-5188

Query Match 13.4%; Score 48.5; DB 1; Length 60;
Best Local Similarity 41.9%; Pred. No. 13;
Matches 13; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 28 PNCNGGICLPGIAGVSGSCPCD-GFTDP 57
DB 28 PTPSSPGSPCSPLLAFHFWSPVCPNAGRTSP 58

RESULT 10
US-09-802-704-17
Sequence 17, Application US/09802704
GENERAL INFORMATION:
APPLICANT: Godbole, Shubhada
APPLICANT: Boyle, Bryan J
APPLICANT: Arterburn, Matthew C
APPLICANT: Liu, Chenghua
APPLICANT: Tang, Y Tom
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AN
TITLE OF INVENTION: POLYNUCLEOTIDES
FILE REFERENCE: HVS-40
CURRENT APPLICATION NUMBER: US/09/802,704
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 17
LENGTH: 36
TYPE: PRT
ORGANISM: Homo sapiens
US-09-802-704-17

Query Match 13.0%; Score 47; DB 5; Length 36;
Best Local Similarity 33.3%; Pred. No. 12;
Matches 9; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

QY 28 PNCNGGICLPGIAGVSGSCPCDGF 54
DB 3 PRLCMHACVTFP----GSYRCPCPGY 25

RESULT 11
US-09-817-427-450
Sequence 450, Application US/09817427
GENERAL INFORMATION:
APPLICANT: Gearing, David P.
APPLICANT: Kingsbury, Gillian A.
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
TITLE OF INVENTION: HUMAN MIXED LYMPHOCYTE LIBRARY
FILE REFERENCE: MLN98-54PA
CURRENT APPLICATION NUMBER: US/09/817,427
CURRENT FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: 09/409,195
PRIOR FILING DATE: 1999-09-30
NUMBER OF SEQ ID NOS: 622
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 450
LENGTH: 60
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(29)
US-09-817-427-450

```
Query Match 12.8%; Score 46.5; DB 5; Length 60;
Best Local Similarity 34.1%; Pred. No. 23;
Matches 15; Conservative 6; Mismatches 8; Indels 15; Gaps 4;

QY 13 LSLGVPQFGKGDICDPNPGENGICLPGLAV-GS--FSCCEPDG 53
DB 28 LAMGSP-----GPGCEDPA---GAGVGRSRRPTCQSPG 59

RESULT 12
US-09-249-697A-11
; Sequence 11, Application US/09249697A
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL
; FILE REFERENCE: 24011-727
; CURRENT APPLICATION NUMBER: US/09/249,697A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Mammalian
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(58)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-09-249-697A-11

Query Match 12.7%; Score 46; DB 5; Length 58;
Best Local Similarity 29.2%; Pred. No. 25;
Matches 7; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 31 CENGICLPGLAVGFSFSCCPDGF 54
DB 14 CLHDGVCMIYEALDKYACNCVGY 37

RESULT 13
US-09-750-972-53
; Sequence 53, Application US/09750972
; GENERAL INFORMATION:
; APPLICANT: Pramod K. Srivastava
; TITLE OF INVENTION: ALPHA(2) MACROGLOBULIN RECEPTOR AS A HEAT SHOCK
; FILE REFERENCE: 8449-134
; CURRENT APPLICATION NUMBER: US/09/750,972
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/750,972
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 09/668,724
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-750-972-53

Query Match 12.3%; Score 44.5; DB 5; Length 40;
Best Local Similarity 35.5%; Pred. No. 26;
Matches 11; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 26 CDP--NPCENG-GICLPGLAVGFSFSCCPDG 53

Query Match 12.8%; Score 44; DB 5; Length 11;
Best Local Similarity 54.5%; Pred. No. 8.7;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 44 GSFSCCPDGF 54
DB 1 GSYRCHCPG 11

Query Match 12.3%; Score 44.5; DB 5; Length 43;
Best Local Similarity 29.3%; Pred. No. 28;
Matches 12; Conservative 7; Mismatches 19; Indels 3; Gaps 2;

QY 22 KGDICDPNP--CENGICLPGLAVGFSFSCCPDGF-TD-PNC 59
DB 1 KACRCQPNENHCLGTETCLVPMRSLCNGVDGSDGSPHC 41

RESULT 15
US-09-802-704-14
; Sequence 14, Application US/09802704
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada
; APPLICANT: Boyle, Bryan J
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Liu, Chenghua
; APPLICANT: Tang, Y Tom
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AN
; FILE REFERENCE: HHS-40
; CURRENT APPLICATION NUMBER: US/09/802,704
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-802-704-14
```

Wed May 23 08:50:33 2001

us-09-237-981-14_copy_33_97.clo.rapn

Page 5

Search completed: May 23, 2001, 08:34:16
Job time: 107 sec

12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:32:28 ; Search time 11.52 Seconds

(without alignments)
108.395 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Perfect score: 362

Sequence: 1 MKRSVAWMLVGLSLGVPQF.....FSCPCPDGFTDPCMSVVEV 65

BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 137559

Minimum DB seq length: 0

Maximum DB seq length: 65

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5a_COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5b_COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6a_COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6b_COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCOTUS.COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	222.5	61.5	57	2	US-08-480-229C-23
2	222.5	61.5	57	2	US-08-659-235C-23
3	127	35.1	25	2	US-08-480-229C-22
4	127	35.1	25	2	US-08-659-235C-22
5	101	27.9	52	3	US-08-978-741-15
6	95	26.2	34	6	5514582-23
7	93.5	25.8	42	2	US-08-480-229C-25
8	93.5	25.8	42	2	US-08-659-235C-25
9	92	25.1	34	6	5514582-28
10	91	25.1	34	6	5514582-26
11	88	24.3	34	6	5514582-27
12	87.5	24.2	43	2	US-08-480-229C-24
13	87.5	24.2	43	2	US-08-659-235C-24
14	86	23.9	34	6	5514582-25
15	82	22.7	34	6	5514582-29
16	77	21.3	34	6	5514582-22
17	77	21.3	47	3	US-08-753-007A-11
18	76	21.0	47	4	US-09-398-496-11
19	75	20.7	36	1	US-08-340-539A-27
20	75	20.7	34	6	5514582-24
21	72	19.9	43	2	US-08-751-305-4
22	72	19.9	49	4	US-09-020-880-18
23	72	19.9	49	4	US-09-020-880-19
24	72	19.9	49	4	US-09-020-880-24
25	72	19.9	49	4	US-09-020-880-26
26	72	19.9	49	4	US-09-020-880-27
27	72	19.9	49	4	US-09-020-880-31

28	71	19.6	36	1	US-08-340-539A-24	Sequence 24, Appl
29	71	19.6	36	1	US-08-340-539A-25	Sequence 25, Appl
30	71	19.6	49	4	US-09-020-880-21	Sequence 21, Appl
31	71	19.6	49	4	US-09-020-880-23	Sequence 23, Appl
32	71	19.6	49	4	US-09-020-880-30	Sequence 30, Appl
33	70	19.3	49	4	US-09-020-880-32	Sequence 32, Appl
34	70	19.3	49	4	US-09-020-880-33	Sequence 33, Appl
35	69	19.1	36	1	US-08-340-539A-26	Sequence 26, Appl
36	69	19.1	45	1	US-08-278-089A-25	Sequence 25, Appl
37	69	19.1	53	1	US-08-284-923-2	Sequence 2, Appl1
38	69	19.1	53	4	US-08-619-032B-2	Sequence 2, Appl1
39	69	19.1	53	6	5332669-2	Sequence 2, Appl1
40	67	18.5	36	1	US-08-340-539A-28	Sequence 28, Appl
41	67	18.5	41	6	5177197-32	Sequence 28, Appl
42	67	18.5	41	6	5177197-40	Sequence 28, Appl
43	67	18.5	49	4	US-09-020-880-14	Sequence 14, Appl
44	67	18.5	49	4	US-09-020-880-16	Sequence 16, Appl
45	66	18.2	42	6	5177197-41	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-08-480-229C-23

Sequence 23, Application US/08480229C

Patent No. 5874562

GENERAL INFORMATION:

APPLICANT: Quartermou, Thomas

APPLICANT: Hogan, Bridgid

APPLICANT: Snodgrass, H. Ralph

APPLICANT: Zupancic, Thomas J.

TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESSES:

ADDRESS: Penite 6 Edmonds LLP

SECRET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn, Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/480,229C

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.

REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-0026-999

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 Penite

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 57 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-480-229C-23

Query Match 61.5% Score 222.5; DB 2; Length 57;

Best Local Similarity 88.9%; Pred. No. 6, 8e-18;

Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 26 CDNPENCGICLPGAVG-----SFCBCEPDGFTDNCSSVVEV 65
Db 1 CDNPENCGICLPGAVGXXXXXSFCBCEPDGFTDNCSSVVEV 45

RESULT 2

US-08-659-235C-23
; Sequence 23, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659/235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-659-235C-23

Query Match 61.5%; Score 222.5; DB 2; Length 57;
Best Local Similarity 88.9%; Pred. No. 6.8e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 26 CDNPENCGICLPGAVG-----SFCBCEPDGFTDNCSSVVEV 65
Db 1 CDNPENCGICLPGAVGXXXXXSFCBCEPDGFTDNCSSVVEV 45

RESULT 3

US-08-480-229C-22
; Sequence 22, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229C-22

Query Match 35.1%; Score 127; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVMLVGLGVPQFGKDI 25
Db 1 MKRSYAVMLVGLGVPQFGKDI 25

RESULT 4
US-08-659-235C-22
; Sequence 22, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659/235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235C-22

Query Match 35.1%; Score 127; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKRSVAWMLVGLSPVPGKGD1 25
DB 1 MKRSVAWMLVGLSPVPGKGD1 25

RESULT 5
US-08-978-741-15
Sequence 15, Application US/08978741
Patent No. 6100076
GENERAL INFORMATION:
APPLICANT: Yang Wang, Michael W. Spellman
TITLE OF INVENTION: O-Fucosyltransferase
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,741
FILING DATE: 26-NO. 6100076-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/792498
FILING DATE: 31
ATTORNEY/AGENT INFORMATION:
NAME: Syvoda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1041P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-978-741-15

Query Match 27.9%; Score 101; DB 3; Length 52;
Best Local Similarity 48.6%; Pred. No. 0.00011;
Matches 18; Conservative 3; Mismatches 14; Indels 2; Gaps 1;
OY 23 GDICDNPENGICLPGLAVGSFSCPCPDGFTDPNC 59
DB 4 GDICDNPENGICLPGLAVGSFSCPCPDGFTDPNC 38

RESULT 6
5514582-23
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 23:
LENGTH: 34
5514582-23

Query Match 26.2%; Score 95; DB 6; Length 34;
Best Local Similarity 38.9%; Pred. No. 0.00032;
Matches 14; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

OY 26 CDPNPENGICLPGLAVGSFSCPCPDGFTDPNCSS 61
DB 1 CASAPQNGVCIDG--VNGYMCDCQPGYTGTHCET 34

RESULT 7
US-08-480-229C-25
Sequence 25, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quetters, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:

Wed May 23 08:50:31 2001

us-09-237-981-14_copy_33_97.clo.ra1

Page 4

LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229C-25

Query Match 25.8%; Score 93.5; DB 2; Length 42;
Best Local Similarity 43.6%; Pred. No. 0.00058;
Matches 17; Conservative 5; Mismatches 12; Indels 5; Gaps 1;

QY 26 CDPNPGNGGICLPGLA-----VGSFSCGCPDGFDPNC 59
DB 1 CEVEPCKNGGICITDVAXXXXXXXNYSCGCPGFGRNC 39

RESULT 8

US-08-659-235C-25
Sequence 25, Application US/08659235C
Patent No. 5877281

GENERAL INFORMATION:
APPLICANT: Queternous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penrite & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
ZIP: 10036-2711
COUNTRY: United States

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Penrite

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235C-25

Query Match 25.8%; Score 93.5; DB 2; Length 42;
Best Local Similarity 43.6%; Pred. No. 0.00058;
Matches 17; Conservative 5; Mismatches 12; Indels 5; Gaps 1;

QY 26 CDPNPGNGGICLPGLA-----VGSFSCGCPDGFDPNC 59
DB 1 CEVEPCKNGGICITDVAXXXXXXXNYSCGCPGFGRNC 39

RESULT 9

5514582-28
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 28
LENGTH: 34
5514582-28

Query Match 25.4%; Score 92; DB 6; Length 34;
Best Local Similarity 50.0%; Pred. No. 0.00067;
Matches 17; Conservative 2; Mismatches 13; Indels 2; Gaps 1;

QY 26 CDPNPGNGGICLPGLAVGSFSCGCPDGFDPNC 59
DB 1 CLENPCNGGICVQHRE--SFSCDPCPGFYNGC 32

RESULT 10
5514582-26
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 26
LENGTH: 34
5514582-26

Query Match 25.1%; Score 91; DB 6; Length 34;
Best Local Similarity 44.1%; Pred. No. 0.00087;
Matches 15; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

QY 26 CDPNPGNGGICLPGLAVGSFSCGCPDGFDPNC 59
DB 1 CETSPCONGCKKDGK--GEYTCCLBSFEGKNC 32

RESULT 11
5514582-27
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
5514582-27

Query Match 25.1%; Score 91; DB 6; Length 34;
Best Local Similarity 44.1%; Pred. No. 0.00087;
Matches 15; Conservative 6; Mismatches 11; Indels 2; Gaps 1;

Wed May 23 08:50:31 2001

us-09-237-981-14_copy_33_97.clo.ra1

Page 6

APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 25
LENGTH: 34
514582-25

Query Match 23.8%; Score 86; DB 6; Length 34;
Best Local Similarity 41.2%; Pred. No. 0.003; Mismatches 12; Indels 2; Gaps 1;
Matches 14; Conservative 6; Mismatches 12; Indels 2; Gaps 1;
OY 26 CDPNPGCNGGICLPGIAGVSFSCPCPDGFTDPNC 59
DB 1 CEGHPCLNGHCKDG-IGDYTCYCAEGEGKNC 32

RESULT 15
514582-29
Patent No. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 29
LENGTH: 34
514582-29

Query Match 22.7%; Score 82; DB 6; Length 34;
Best Local Similarity 47.1%; Pred. No. 0.0083; Mismatches 14; Indels 2; Gaps 1;
Matches 16; Conservative 2; Mismatches 14; Indels 2; Gaps 1;
OY 26 CDPNPGCNGGICLPGIAGVSFSCPCPDGFTDPNC 59
DB 1 CRTNPLHGRCLE--VEGHRLCCHCPVGYGPRC 32

Search completed: May 23, 2001, 08:33:03
Job time: 35 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:34:19 ; Search time 18.56 Seconds
(without alignments)
1503.314 Million cell updates/sec

Title: us-09-237-981-14_Copy_108_513

Sequence: 1 GPCPNCHNGTCEISEAY.....MSWGRITLASELLGTEEE 406

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 142050

Minimum DB seq length: 0
Maximum DB seq length: 406

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1072	47.3	401	2	glycoprotein anti g
2	643	28.4	218	2	milk fat globule p
3	464.5	20.5	216	2	factor VIII-associ
4	209.5	9.3	293	2	neurogenic repetit
5	197.5	8.7	387	2	Meich protein
6	189.5	8.4	308	3	epidermal growth f
7	187	8.3	383	2	delta-like homola
8	179	7.9	385	2	homocysteine beta
9	179	7.9	385	2	predilec protei
10	176.5	7.8	252	2	chondroitin sulfat
11	169.5	7.5	102	2	hypothetical prote
12	140.5	6.2	379	2	hypothetical prote
13	134.5	5.8	149	1	discoildin I chain
14	134.5	5.8	149	1	discoildin I chain
15	134.5	5.8	149	1	discoildin I chain
16	129.5	5.7	253	2	discoildin I chain
17	129.5	5.7	253	2	discoildin I chain
18	129.5	5.7	253	2	discoildin I chain
19	129.5	5.7	253	2	discoildin I chain
20	129.5	5.7	253	2	discoildin I chain
21	129.5	5.7	253	2	discoildin I chain
22	129.5	5.7	253	2	discoildin I chain
23	129.5	5.7	253	2	discoildin I chain
24	129.5	5.7	253	2	discoildin I chain
25	129.5	5.7	253	2	discoildin I chain
26	129.5	5.7	253	2	discoildin I chain
27	129.5	5.7	253	2	discoildin I chain
28	129.5	5.7	253	2	discoildin I chain
29	129.5	5.7	253	2	discoildin I chain

30	93	4.1	356	2	T20656	hypothetical prote
31	92.5	4.1	402	2	S42367	lag-2 protein - Ca
32	92	4.1	302	2	T22899	hypothetical prote
33	91	4.0	264	2	T16271	hypothetical prote
34	90.5	4.0	399	2	T02678	hypothetical prote
35	90	4.0	356	2	A25918	thrombomodulin - b
36	90	4.0	372	2	JC5377	I-selectin precurs
37	89.5	4.0	390	2	DB2922	conserved hypothet
38	89	3.9	171	2	149612	gene cripo protei
39	88	3.9	291	2	I38098	t-plasminogen acti
40	87.5	3.9	349	2	S62137	SEC13 protein homo
41	87	3.8	281	2	S39495	u-plasminogen acti
42	87	3.8	331	2	DB1409	probable periplasm
43	87	3.8	370	2	S22124	L-selectin precurs
44	86.5	3.8	314	2	T28879	hypothetical prote
45	86.5	3.8	370	2	C82553	homoserine O-acety

ALIGNMENTS

RESULT 1
655138
glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
N:Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E
C:Species: Bos primigenius taurus (cattle)
C>Date: 28-Oct-1996 #sequence_Revision 13-Mar-1997 #text_Change 07-Aug-1998
C:Accession: S65138; G48394
R:Ack: N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A>Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal
A:Reference number: S65138; MUID:96125736
A:Accession: S65138
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-401 <AOK>
R:Method: I. H.; Banghart, L. R.; Lane, W. S.
Biochem. Mol. Biol. Int. 28, 545-554, 1993
A>Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-
IT-like sequences.
A:Reference number: A48394; MUID:93250576
A:Accession: G48394
A>Status: preliminary
A:Molecule type: protein
A:Residues: 207-220 <MAT>
A:Experimental source: milk
A>Note: sequence extracted from NCBI backbone (NCBI:131457)
C:Superfamily: milk fat globule protein; discolidin I amino-terminal homology; EGF hom
C:Keywords: glycoprotein
F:1-32/Domain: EGF homology <EG3>
F:40-76/Domain: EGF homology <EG3>
F:82-219/Domain: discolidin I amino-terminal homology <DN1>
F:243-401/Domain: discolidin I amino-terminal homology <DN2>

Query Match 47.3% Score 1072; DB 2; Length 401;
Best Local Similarity 47.9% Pred. No. 3.1e-74;
Matches 193; Conservative 68; Mismatches 102; Indels 40; Gaps 3;
OY 1 GPCPNCHNGTCEIS- EAYKDPFIYGVCKRCRNGICGCHINCEVEPCKNGIC 59
DB 38 GPCPNCHNGTCEIS- EAYKDPFIYGVCKRCRNGICGCHINCEVEPCKNGIC 59
OY 60 TDIVANVSCGEPGFMGRNCQKCSPIGEGILTSMOGTASSYHRLFLQAWPYRYA 119
DB 60 TDIVANVSCGEPGFMGRNCQKCSPIGEGILTSMOGTASSYHRLFLQAWPYRYA 119
OY 83 -----CTSPAGQGTALDSQISMSHGLGFMALGFMALDEL 119
DB 120 PLANKGLINAMTAENDRMKRWIOINLQKRRVGYTGQAKRISPEYIFKRYASND 179
OY 120 PLANKGLINAMTAENDRMKRWIOINLQKRRVGYTGQAKRISPEYIFKRYASND 179
DB 120 PLANKGLINAMTAENDRMKRWIOINLQKRRVGYTGQAKRISPEYIFKRYASND 179
OY 180 GKTAMTVYKKTNDKVFKNQNDNNTPTANSFTPKIAQYVRLYQVCRKCTIRMEILG 239
DB 180 GKTAMTVYKKTNDKVFKNQNDNNTPTANSFTPKIAQYVRLYQVCRKCTIRMEILG 239

DB 179 GROFOFOYVAGSGDKIFPGVNNNSGLKINLEDFPLETOYVRLVPLICHRGCTLRFELG 238

OY 240 CELSGSEFPLGKMSGHIDQYOTASSIFPTLMDMFEMPERKRLDKOGKXVNAWTSCHND 299

DB 239 CELNGCTEPDLGKMDTIPNKQITAYSSYKWTGLSFAFMPFYARLDNGCKFNAMTAQTN 298

OY 300 QSOVLQVLLVPTKYTGIIIOGAKDGHQFPGSYKILAYSNDEHMTYODEKQKRXV 359

DB 299 ASEMLDIDGSGKRYTGITIOGARDFGHIOYAAVRAVAGDDGVMTWTEKPGASEKIF 358

OY 360 QGNFNDTRKKNVIDPPIYARHRIPLPMSWGRITLASSELLG 402

DB 359 PGNNNNHSHKKNIFETPQARFVRIOYVAMHNRITLVELLG 401

RESULT 2

A47285
milk fat globule protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 28-May-1999
C:Accession: A47285
R:Laococa, D.; Peterson, J.A.; Ureia, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
Cancer Res. 51, 4994-4998, 1991
A>Title: A M-r-46,000 human milk fat globule protein that is highly expressed in human B
A:Reference number: A47285; MUID:91371351
A:Accession: A47285
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <LARG>
A:Cross-references: GB:56151; NID:9235396; PIDN:AA819771.1; PID:9235397
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
F.1-56/Domain: discoidin I amino-terminal homology (fragment) <N1>
F.60-218/Domain: discoidin I amino-terminal homology <DN2>

Query Match 28.4% Score 643; DB 2; Length 218;
Best Local Similarity 56.8%; Pred. No. 6,6e-42;
Matches 117; Conservative 30; Mismatches 59; Indels 0; Gaps 0;

OY 197 FRGJINNTFVANSFPTPIKAOYVRLVPOVCRHCTRLMELGCELSCGSEPLDMKSGH 256

DB 13 FVGMKMKMNVHNLFEFPEVNOYVRLVPTSCHTCTLRFLGCELNCANPLDKMNSI 72

OY 257 ODYOTASSIFPTLMDMFEMPERKRLDKOGKXVNAWTSCHNDQSOVLQVLLVPTKY 316

DB 73 PDQITASSSVKWTGLHLEFEMNSYARLDKQGNNAVASYGNQWLQVLDGSSKEVIG 132

OY 317 IITOGAKDGHQFPGSYKILAYSNDEHMTYODEKQKRXVQGNFNDTRKKNVIDP 376

DB 133 IITOGAKDGHQFPGSYKILAYSNDEHMTYODEKQKRXVQGNFNDTRKKNVIDP 376

OY 377 IYARHRIPLPMSWGRITLASSELLG 402

DB 193 ILARYRILPVMHNRITLVELLG 401

RESULT 3

A44258
factor VIII-associated gene B hypothetical protein - human
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
C:Accession: A44258
R:Levinson, B.; Kenwick, S.; Gamel, P.; Fisher, K.; Glitschier, J.
Genomics 14, 585-589, 1992
A>Title: Evidence for a third transcript from the human factor VIII gene.
A:Reference number: A44258; MUID:93052386
A:Accession: A44258
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-216 <LEV>
A:Cross-references: GB:M90707; NID:9182316; PIDN:AAA58466.1; PID:9182317
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F.1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>

F:57-210/Domain: discoidin I amino-terminal homology <DN2>

Query Match 20.5% Score 464.5; DB 2; Length 216;
Best Local Similarity 46.7%; Pred. No. 2.5e-28;
Matches 99; Conservative 31; Mismatches 75; Indels 7; Gaps 3;

OY 196 VFGNIDNNTFVANSFPTPIKAOYVRLVPOVCRHCTRLMELGCELSCGSEPLDMKSGH 255

DB 9 VFGNIDNNTFVANSFPTPIKAOYVRLVPOVCRHCTRLMELGCELSCGSEPLDMKSGH 255

OY 256 IDYOTASSIFPTLMDMFEMPERKRLDKOGKXVNAWTSCHNDQSOVLQVLLVPTKY 314

DB 69 ISDAQITASSYF---TNMFATWSPSKARLHLQGRSMARPPQVNNPKEMLDVDFOKTKV 124

OY 315 TGIITOGAKDGHQFPGSYKILAYSNDEHMTYODEKQKRXVQGNFNDTRKKNVID 374

DB 125 TGVITOGAKDGHQFPGSYKILAYSNDEHMTYODEKQKRXVQGNFNDTRKKNVID 374

OY 375 PPIYARHRIPLPMSWGRITLASSELLG 406

DB 183 PPIYARHRIPLPMSWGRITLASSELLG 406

RESULT 4

B26637
neurogenic repetitive locus 95F protein - fruit fly (Drosophila melanogaster) (fragme
C:Species: Drosophila melanogaster
C>Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 21-Jul-2000
C:Accession: B26637
R:Kunst, E.; Dietrich, U.; Teepas, U.; Bremner, K.A.; Weigel, D.; Vaesslin, H.; Campos-
Embo, J. 6, 761-766, 1987
A>Title: EGF homologous sequences encoded in the genome of Drosophila melanogaster. a
A:Reference number: A91081; MUID:87218537
A:Accession: B26637
A:Molecule type: mRNA
A:Residues: 1-293 <RNU>
A:Cross-references: GB:X05144; NID:97519; PIDN:CAA28793.1; PID:9295356
C:Gene: flybase:cyb
A:Cross-references: Flybase:FBgn0000368
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: transmembrane protein; EGF1
F.139-170/Domain: EGF homology <EGF1>
F.177-208/Domain: EGF homology <EGF1>
F.216-252/Domain: EGF homology <EGF>

Query Match 9.3% Score 209.5; DB 2; Length 293;
Best Local Similarity 45.5%; Pred. No. 9.1e-09;
Matches 53; Conservative 12; Mismatches 23; Indels 7; Gaps 2;

OY 3 CTIPNCHNGCTCELSAAYKGFDFIYVCKPGRFGNIRHONHINECEVERCKNGGICTL 62

DB 139 CLNTECLNNGTC-TNOY-----AAFFQCPGPGGGRHCEONIDECADQPCFNGKCTIL 191

OY 63 VANYSCGCPGPGGRNC 79

DB 192 IASIVCDCEPDYKQPC 208

RESULT 5

B49175
Notch A protein - mouse (fragment)
N:Alternate names: Notch homolog
C:Species: Mus musculus (house mouse)
C>Date: 21-Jan-1994 #sequence_revision 05-Jan-1996 #text_change 20-Sep-1999
R:Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A>Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety
A:Reference number: A49175; MUID:93178563
A:Accession: B49175

Db 49 AMCSSIVDTNQYIVAGCEVPRTEMCVALQGRD--HDQWTSYKIRYSLDNVTWSEY--- 103
QY 352 KQKDKYFQGNFENDTHRKNVIDPPIYARHRIILPWSWYGRITLASE 398
Db 104 --RNGAALITGVTDNRTVYNHFFDTPIRARSIAIHPLTWNHNSISLCE 148

Search completed: May 23, 2001, 08:35:56
Job time: 97 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:35:39 ; Search time 11.79 seconds

(without alignments)
1179.621 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Perfect score: 2264
Sequence: 1 GPCPNPCNHCNGTCETSEAY.....MSWYGRITLASELLGCTEE 406

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 63303

Minimum DB seq length: 0

Maximum DB seq length: 406

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020.5	45.1	387	1	MFQM_HUMAN
2	295.5	13.1	280	1	Q9W6T3_tugu_rubrip
3	280	12.4	224	1	XLR1_MOUSE
4	267	11.8	224	1	XLR1_HUMAN
5	187	8.3	383	1	DLK_HUMAN
6	179	7.9	385	1	DLK_MOUSE
7	141.5	6.2	379	1	WIF1_MOUSE
8	140.5	6.2	257	1	DIS2_DICDI
9	140.5	6.2	379	1	WIF1_HUMAN
10	134.5	5.9	149	1	DISC_DICDI
11	134.5	5.9	253	1	DISC_DICDI
12	122.5	5.4	396	1	PRP2_BOVIN
13	121.5	5.4	253	1	DISA_DICDI
14	117	5.2	400	1	PRP2_HUMAN
15	99.5	4.4	376	1	FA10_TROCA
16	94	4.2	188	1	CR11_HUMAN
17	92.5	4.1	402	1	LAG2_CABEL
18	90	4.0	356	1	TRBM_BOVIN
19	90	4.0	372	1	LEMI_MACU
20	90	4.0	372	1	LEMI_PAPPA
21	89.5	4.0	58	1	MT_POTPO
22	89	3.9	171	1	CR10_MOUSE
23	88	3.9	188	1	CR12_HUMAN
24	87.5	3.9	349	1	SEH1_YEAST
25	87	3.8	370	1	LEMI_BOVIN
26	86.5	3.8	370	1	METX_XYLEA
27	86	3.8	230	1	SPTT_DROME
28	85.5	3.8	402	1	RL3_HUMAN
29	85	3.7	255	1	LEMI_HUMAN
30	84.5	3.7	372	1	LEMI_HUMAN
31	84.5	3.7	372	1	LEMI_PANTR
32	84	3.7	372	1	LEMI_PONRY
33	84	3.7	64	1	MTCD_HELPO

34	83	3.7	273	1	TYRO_CORUA	Q08410_coturnix_co
35	83	3.7	318	1	CH11_SOLRU	P52403_solanum_tub
36	82.5	3.6	119	1	ANTA_HAECH	P16242_haemantaria
37	82.5	3.6	335	1	COID_HUMAN	P15813_homo_sapien
38	82.5	3.6	372	1	LEMI_MOUSE	P18337_mus_musculu
39	81	3.6	171	1	LEMI_HYDAT	O27262_hydra_atten
40	81	3.6	318	1	CH13_SOLRU	P52405_solanum_tub
41	81	3.6	335	1	EDSP_CINPY	P46058_cynops_pyrt
42	80.5	3.6	372	1	LEMI_RAT	P30836_rattus_norv
43	80	3.5	348	1	Y217_AQUAE	O65408_aquilex_aeo
44	80	3.5	372	1	AG1_DRTDI	P11218_urtica_dio1
45	79	3.5	74	1	MT_CRAVI	P23038_crassostrea

ALIGNMENTS

RESULT	ID	MEQM_HUMAN	STANDARD	PRT	387 AA.
AC	008431				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGG FACTOR 8) (MFG-EG) (HMF6)				
DE	(BREST EPITHELIAL ANTIGEN BA46) (MEQM) [CONTAINS: MEDLN].				
GN	MEQB8.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Breast; and Breast carcinoma;				
RX	MEDLINE=96213908; PubMed=8639264;				
RA	Conto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Peterson J.A.;				
RT	"Cloning and sequence analysis of human breast epithelial antigen				
RT	BA46 reveals an RGD cell adhesion sequence presented on an epidermal				
RT	growth factor-like domain.";				
RL	DNA Cell Biol. 15:281-286(1996).				
RN	[2]				
RP	SEQUENCE OF 170-387 FROM N.A.				
RC	TISSUE=Mammary gland;				
RX	MEDLINE=91371351; PubMed=1909932;				
RA	Larocca D., Peterson J.A., Urra R., Kunyoshi J., Bistrain A.M.,				
RA	Ceriani R.L.;				
RT	"A Mr 46,000 human milk fat globule protein that is highly expressed				
RT	in human breast tumors contains factor VIII-like domains.";				
RL	Cancer Res. 51:4994-4998(1991).				
RN	[3]				
RP	PARTIAL SEQUENCE, AND CHARACTERIZATION.				
RC	TISSUE=Milk;				
RX	MEDLINE=98194924; PubMed=9535276;				
RA	Guiffrida M.G., Cavaleto M., Giunta C., Conti A.,				
RA	Godovac-Zimmermann J.;				
RT	"Isolation and characterization of full and truncated forms of human				
RT	breast carcinoma protein BA46 from human milk fat globule membranes.";				
RL	J. Protein Chem. 17:143-148(1998).				
RN	[4]				
RP	SEQUENCE OF 268-317, AND IDENTIFICATION OF MEDLN.				
RX	MEDLINE=95342076; PubMed=10411933;				
RA	Heesqvist B., Neslund J., Stelten K., Westermark G.T., Mucchiano G.,				
RA	Tjernberg L.O., Norstedt C., Engstrom U., Westermark P.;				
RT	"Medin: an integral fragment of aortic smooth muscle cell-produced				
RT	lactadherin forms the most common human amyloid.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 96:8669-8674(1999).				
RN	[5]				
RP	CHARACTERIZATION.				
RX	MEDLINE=97405885; PubMed=9260929;				
RA	Taylor M.R., Conto J.R., Scallan C.D., Ceriani R.L., Peterson J.A.;				
RA	"Lactadherin (formerly BA46), a membrane-associated glycoprotein				
RT	expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp				
RT	(RGD)-dependent cell adhesion.";				

RL DNA Cell Biol. 16:861-869(1997).
CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. BINDS
CC SPECIFICALLY TO ROTAVIRUS AND INHIBITS ITS REPLICATION.
CC -1- FUNCTION: MEDIN IS THE MAIN CONSTITUENT OF AORTIC MEDIAL AMYLOID.
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND AORTIC
CC MEDIA. OVEREXPRESSED IN SEVERAL CARCINOMAS.
CC -1- PTM: MEDIN HAS A RAGED N-TERMINUS WITH MINOR SPECIES STARTING AT
CC AMINO ACID 264 AND 273.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U58516; AAC50549.1; -
CC EMBL, S56151; AAB19771.1; -
CC MIM: 602281; -
CC InterPro: IPR000421; -
CC InterPro: IPR000561; -
CC Pfam: PF00008; EGF_1.
CC Pfam: PF00754; F5_F8_type_C; 2.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC PROSITE: PS01285; FAS58C_1; 2.
CC PROSITE: PS01286; FAS58C_2; 2.
CC Signal: Glycoprotein; Milk; Repeat; EGF-like domain; Amyloid.
FT SIGNAL 1 387
FT CHAIN 24 387 LACTADHERIN.
FT CHAIN 202 387 LACTADHERIN; SHORT FORM.
FT DOMAIN 266 317 MEDIN.
FT DOMAIN 24 67 EGF-LIKE.
FT DOMAIN 70 225 F5/8 TYPE C 1.
FT DOMAIN 230 387 F5/8 TYPE C 2.
FT SITE 46 48 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 27 38 BY SIMILARITY.
FT DISULFID 32 55 BY SIMILARITY.
FT DISULFID 57 66 BY SIMILARITY.
FT DISULFID 70 225 BY SIMILARITY.
FT DISULFID 212 216 BY SIMILARITY.
FT CARBOHYD 238 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 387 AA; 43123 MW; 2EE6571DEC83782D CRC64;

Query Match 45.1%; Score 1020.5; DB 1; Length 387;
Best Local Similarity 48.8%; Pred. No. 8.9e-74;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;

QY 3 CTPNCHNGTTC-EISBAPRGDTFTGYCKCPREFNCHGCHNHNCEVEVECKNGICTD 61
DB 27 CSKNCHNGGCEISIQEVRGDPFSPYCTCLKAYA-- 65
QY 62 LVANTSCBCEPDEFGNRCQKSGPLGIEGGIISNOQITASSSTRALFGLQKWPYVARL 121
DB 66 ----- 111
QY 122 NKGKLNMTAEADRMKRMIOINLQRMKRYTGVITGARGISPEYIKFYAIANSNGK 181
DB 109 NAGGVNANMPSSND- NPMIOVNLRRMWTGVVYGASRLASHHYLAFAVVASLNGH 167
QY 182 TW-AMYKAKGTNEMVFRGINDNTPYANSTPPIKQAVRYLAPQVCRHCHLMBELIIGC 240
DB 166 EEDFTHDYNKHKFEV--GMKKNVAVNLEFTEVEQVRYRLPTFSCHACTLARELLIGC 225

QY 241 ELSCGSEPLGKSGCHIDVOYTASSIFRTLMKMEFWEPKRALDKOGKVAMTSGHNDQ 300
DB 226 ELINGANLEPLGKSNKSPDKOATASSSYKTWGLHLEFSPNPAYALDKOGKFNAWVAGSYGN 285
QY 301 SOMLOVDLVPYTKVGTITQGKDFGQVFGSYKLVASNDGEMHTVYODEKORKDYVQ 360
DB 286 DQMLQVDGSSKEVGTITQGARNGSQVFAVASYKVASNDASAMMETEYODPRTGSKIFP 345
QY 361 GNFNDPFRKRVIDPPIYARHRIILPMSMGRITLAEELGCG 402
DB 346 GMDNHSKKNLFEETPLARYRLPLVAMHRIALRELLGCG 387

RESULT 2
XLR1_FUGRU STANDARD: PRT: 280 AA.
ID XLR1_FUGRU
AC Q9W0R5
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE X-LINKED JUVENILE RETINOSCHISIS PROTEIN PRECURSOR.
GN XLR1.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99263230; PubMed=10330123;
RA Brunner B., Todt T., Lenzner S., Stout K., Schulz U., Ropers H.-H.,
RA Kalschauer V.M.;
RT "Genomic structure and comparative analysis of nine Fugu genes:
RT conservation of synteny with human chromosome Xp22.2-p22.1.";
RL Genome Res. 9:437-448(1999).
CC -1- FUNCTION: MAY BE ACTIVE IN CELL ADHESION PROCESSES DURING RETINAL
CC DEVELOPMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF146687; AAD28797.1; -
CC InterPro: IPR000421; -
CC Pfam: PF00754; F5_F8_type_C; 1.
CC PROSITE: PS01285; FAS58C_1; 1.
CC PROSITE: PS01286; FAS58C_2; FALSE-NEG.
CC Cell adhesion; Signal.
FT SIGNAL 1 280
FT CHAIN 119 275 X-LINKED JUVENILE RETINOSCHISIS PROTEIN.
FT DOMAIN 119 275 F5/8 TYPE C.
FT DISULFID 119 275 BY SIMILARITY.
SQ SEQUENCE 280 AA; 32032 MW; ADF9A3222ED0167 CRC64;

Query Match 13.1%; Score 295.5; DB 1; Length 280;
Best Local Similarity 42.2%; Pred. No. 1.7e-16;
Matches 68; Conservative 24; Mismatches 54; Indels 15; Gaps 5;

QY 248 PLGMSGHIQYDVTASSIFRTLMKMEFWEPKRALDKOGKVAMTSGHNDQ 302
DB 124 PLGFAGSISPDQITCS-----NDQYTAFFSSWLPSSKARLNTQFGCAWLKRDQNTQ 177
QY 303 KIQVQLLPTVTVVGIITQGKDFGQVFGSYKLVASNDGEMHTVYODEKORKDYVQ 361
DB 178 WLOIDLIAKVVSIILQGRCDAD--EWITVYSLOQYRTDKLWMLTYKQD--IGNRVYVYG 234

```

OY 362 NFDNDTRKNVDPPIYARHRLPMSWYGTILASLLOC 402
DB 235 NSDRSSVONLPRPIVAKIRILPLGWHRIALRLELLC 275

RESULT 3
XLR1_MOUSE STANDARD: PRT: 224 AA.
ID XLR1_MOUSE
AC Q92114;
DT 30-MAY-2000 (rel. 39, Created)
DR 30-MAY-2000 (rel. 39, Last sequence update)
DE 30-MAY-2000 (rel. 39, Last annotation update)
DE X-LINKED JUVENILE RETINOSCHISIS PROTEIN PRECURSOR.
GN RSI OR RSH OR XLR1.
OS Mus musculus (mouse); Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Murielae; Mus.
OX MGI:Taxid=10990;
PB 1;
RC SEQUENCE FROM N.A.
RA MEDLINE=99160474; PubMed=10051329;
RA Gehrig A.E., Warneke-Wittstock R., Sauer C.G., Weber B.H.F.;
RA "Genomic structure and comparative analysis of seven contiguous genes
RA p22.1";
RA Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RN TISSUE=Retina;
RN MEDLINE=99148018; PubMed=10023077;
RN Reid S.N., Akhmedov N.B., Piltsev N.I., Kozak C.A., Danciger M.,
RN Farber D.B.;
RN "Type mouse X-linked juvenile retinoschisis cDNA: expression in
RN photoreceptors";
RN Gene 227:257-266(1999).
RL -1- PUNCTION: MAY BE ACTIVE IN CELL ADHESION PROCESSES DURING RETINAL
RL DEVELOPMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: RETINA-SPECIFIC.
CC -1- SIMILARITY: CONTAINS 1 FS/8 TYPE C DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: AF084561; AAD21808.1; -
DR EMBL: AF084567; AAD21809.1; -
DR EMBL: AF084562; AAD21809.1; JOINED.
DR EMBL: AF084563; AAD21809.1; JOINED.
DR EMBL: AF084564; AAD21809.1; JOINED.
DR EMBL: AF084565; AAD21809.1; JOINED.
DR EMBL: AF084566; AAD21809.1; JOINED.
DR EMBL: AJ011381; CAA09601.1; -
DR MGD: MG1:1336189; Rsh.
DR InterPro: IPR000421; -
DR Pfam: PF00754; F5_F8_Type_C; 1.
DR PROSITE: PS01285; FA58C.1; 1.
DR PROSITE: PS01286; FA58C.2; FALSE_NCG.
KW Cell adhesion, signal.

```

```

FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 224 X-LINKED JUVENILE RETINOSCHISIS PROTEIN.
FT DOMAIN 63 216 FS/8 TYPE C.
FT DISUPID 63 219 BY SIMILARITY.
SQ SEQUENCE 224 AA; 25575 MW; 4536203CC00E90E4 CRC64;

Query Match 12.4%; Score 280; DB 1; Length 224;
Best Local Similarity 33.8%; Pred. No. 2,2e-15;
Matches 67; Conservative 33; Mismatches 74; Indels 24; Gaps 7;

OY 223 YPOVGRHCHTLMELL--GCELSGC-----SPLMGKSGHIDQYQTASSIFRTIMND 273
DB 34 YQACKCDQOVGAMNLSAGATSLDCIPCPYKRPDLFESGEVTPQITCS-----NPE 87
OY 274 MF-----TWEPKARLDKQGVANMTSGHNSQWLDVDTVPTKVTGTTIGAKDFGHV 328
DB 88 QYGVWYSWTANKARLNSGCGCAMLKYQSSQWLDIDKEIKYISGILTIQGRCDID-- 145
OY 329 QPVGSKYKLVANDGE-HWTVYODEKQKDVFOGDFNDTRHKNVDPPIYARHRLPLW 387
DB 146 EWTYKTSVQYFRDERLNMVYKDO-TGNRRVPYGNDRSSVONLPRPIISRILPL 204
OY 388 SWGRITLASLGGTEE 405
DB 205 GWHYRIAIRMELLCASK 222

RESULT 4
XLR1_HUMAN STANDARD: PRT: 224 AA.
ID XLR1_HUMAN
AC O15537;
DT 30-MAY-2000 (rel. 39, Created)
DR 30-MAY-2000 (rel. 39, Last sequence update)
DE 01-OCT-2000 (rel. 40, Last annotation update)
DE X-LINKED JUVENILE RETINOSCHISIS PROTEIN PRECURSOR.
GN RSI OR XLR1.
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI:Taxid=9606;
PB 1;
RC SEQUENCE FROM N.A. AND VARIANTS XLR1 ARG-96 AND TRP-102.
RA TISSUE=Retina;
RA MEDLINE=97467726; PubMed=9326935;
RA Sauer C.G., Gehrig A., Warneke-Wittstock R., Marguardt A., Ewing C.C.,
RA Gibson A., Lorenz B., Jurkles B., Weber B.H.;
RA "Positional cloning of the gene associated with x-linked juvenile
RA retinoschisis";
RA Nat. Genet. 17:164-170(1997).
RN [2]
RN SEQUENCE OF 19-224 FROM N.A.
RN Buid C.;
RN Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RN VARIANTS XLR1 LYS-72 AND LEU-193.
RN MEDLINE=98430995; PubMed=9760195;
RN Hotta Y., Fujiki K., Hayakawa M., Ohta T., Fujimaki T., Tamaki K.,
RN Yokoyama T., Kanai A., Hirakata A., Hida T., Nishina S., Azuma N.;
RN "Japanese juvenile retinoschisis is caused by mutations of the XLR1
RN gene";
RN Hum. Genet. 103:142-144(1998).
RN [4]
RN VARIANT XLR1 CYS-197.
RN MEDLINE=99180499; PubMed=10079181;
RN Shastry B.S., Hejtmancik F.J., Trese M.T.;
RN "Recurrent missense (R197C) and nonsense (Y89X) mutations in the XLR1
RN gene in families with x-linked retinoschisis";
RN Biochem. Biophys. Res. Commun. 256:317-319(1999).
RN [5]
RN VARIANTS XLR1 E-98; C-108; W-109; C-141; K-146; C-200 AND K-215.
RN MEDLINE=99378073; PubMed=10450864;
RN Gehrig A., White K., Lorenz B., Andrassl M., Clemens S., Weber B.H.;

```

"Assessment of RSI in X-linked juvenile retinoschisis and sporadic senile retinoschisis.";
 Clin. Genet. 55:461-465(1999).
 [6]
 VARIANTS XLRSL LYS-72; VAL-74 AND ARG-109.
 MEDLINE-9250773; PubMed-10234514;
 Huopaniemi L., Rantala A., Forsius H., Somer M., de la Chapelle A.,
 Allitalo T.,
 "Three widespread founder mutations contribute to high incidence of
 X-linked juvenile retinoschisis in Finland.";
 Eur. J. Hum. Genet. 7:368-376(1999).
 [7]
 VARIANT XLRSL SER-193.
 Duval P.-A., Marliens F., Griffon J.-M., Millet P., Arnaud B.,
 Hamel C.P.,
 "X-linked retinoschisis with a novel substitutive amino acid (P193S)
 in XLRSL.";
 Hum. Mutat. 13:259-259(1999).
 [8]
 VARIANTS XLRSL K-72; C-89; E-109; C-182 AND L-203.
 MEDLINE-9235559; PubMed-1020153;
 Keshima Y., Shimoda K., Ishida S., Ozawa Y., Kudoh J., Iwata T.,
 Oyuchi T., Shimizu N.,
 "Identification of four novel mutations of the XLRSL gene in Japanese
 patients with X-linked juvenile retinoschisis.";
 Hum. Mutat. 13:338-338(1999).
 [9]
 VARIANTS XLRSL PRO-13 AND ARG-223.
 MEDLINE-20004542; PubMed-1033068;
 Hiriyama K.T., Bingham E.L., Tashar B.M., Ayagari R., Fishman G.,
 Small K.W., Weinberg D.V., Weleber R.G., Lewis R.A., Andreasson S.,
 "Novel mutations in XLRSL causing retinoschisis, including first
 evidence of a 19423-427(1998) sequence change.";
 Hum. Mutat. 13:423-427(1998).
 HUN FUNCTION MAY BE ACTIVE IN CELL ADHESION PROCESSES DURING RETINAL
 DEVELOPMENT.
 -1- SUPRECELLULAR LOCATION: SECRETED.
 -1- TISSUE SPECIFICITY: RETINA-SPECIFIC.
 -1- DISEASE: DEFECTS IN RSI ARE THE CAUSE OF X-LINKED JUVENILE
 RETINOSCHISIS (RS OR XLRSL). IT IS A RECESSIVELY INHERITED VITREO-
 RETINAL DEGENERATION CHARACTERIZED BY MACULAR PATHOLOGY AND BY
 SPLITTING OF THE SUPERFICIAL LAYER OF THE RETINA.
 -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
 -1- DATABASE: NAME-Mutations of the RSI gene;
 NOTE-Retina International's Scientific Newsletter;
 WWW="http://www.ihpa.org/sci-news/xlrsmt.htm".

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC
 DR EMBL; AF018963; AAC18405.1; JOINED.
 DR EMBL; AF018958; AAC18405.1; JOINED.
 DR EMBL; AF018959; AAC18405.1; JOINED.
 DR EMBL; AF018960; AAC18405.1; JOINED.
 DR EMBL; AF018961; AAC18405.1; JOINED.
 DR EMBL; AF018962; AAC18405.1; JOINED.
 DR EMBL; AF014459; AAC17928.1; JOINED.
 DR EMBL; Z92542; CAB38871.1; JOINED.
 DR MIM; 312700;
 DR Interpro; IPR000421;
 DR Pfam; PF00754; F5_F8_Type_C_1.
 DR PROSITE; PS01285; FAS5C_1; 1.
 DR PROSITE; PS01286; FAS5C_2; FALSE_NEG.
 DR Cell adhesion; Signal; Vision; Disease mutation: Polymorphism.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 224 X-LINKED JUVENILE RETINOSCHISIS PROTEIN.
 FT DOMAIN 63 216 F5/8 TYPE C.

FT DISULFID 63 219
 FT VARIANT 12 12
 FT VARIANT 13 13
 FT VARIANT 59 59
 FT VARIANT 65 65
 FT VARIANT 70 70
 FT VARIANT 70 70
 FT VARIANT 72 72
 FT VARIANT 72 72
 FT VARIANT 74 74
 FT VARIANT 89 89
 FT VARIANT 96 96
 FT VARIANT 98 98
 FT VARIANT 102 102
 FT VARIANT 102 102
 FT VARIANT 103 103
 FT VARIANT 108 108
 FT VARIANT 109 109
 FT VARIANT 109 109
 FT VARIANT 109 109
 FT VARIANT 109 109
 FT VARIANT 110 110
 FT VARIANT 112 112
 FT VARIANT 113 113
 FT VARIANT 127 127
 FT VARIANT 135 135
 FT VARIANT 136 136
 FT VARIANT 138 138
 FT VARIANT 140 140
 FT VARIANT 140 140
 FT VARIANT 141 141
 FT VARIANT 141 141
 FT VARIANT 141 141
 FT VARIANT 141 141
 FT VARIANT 142 142
 FT VARIANT 143 143
 FT VARIANT 146 146
 FT VARIANT 146 146
 FT VARIANT 155 155
 FT VARIANT 155 155
 BY SIMILARITY.
 L -> H (IN XLRSL).
 /FTID-VAR_008209.
 L -> P (IN XLRSL).
 /FTID-VAR_008210.
 C -> S (IN XLRSL).
 /FTID-VAR_008211.
 Y -> C (IN XLRSL).
 /FTID-VAR_008212.
 G -> A (IN XLRSL).
 /FTID-VAR_008213.
 G -> S (IN XLRSL).
 /FTID-VAR_008214.
 E -> D (IN XLRSL).
 /FTID-VAR_008180.
 E -> K (IN XLRSL).
 /FTID-VAR_008181.
 G -> V (IN XLRSL).
 /FTID-VAR_008182.
 Y -> C (IN XLRSL).
 /FTID-VAR_008215.
 W -> R (IN XLRSL).
 /FTID-VAR_008183.
 A -> E (IN XLRSL).
 /FTID-VAR_008216.
 R -> Q (IN XLRSL).
 /FTID-VAR_008217.
 R -> W (IN XLRSL).
 /FTID-VAR_008184.
 L -> R (IN XLRSL).
 /FTID-VAR_008218.
 Y -> C (IN XLRSL).
 /FTID-VAR_008219.
 G -> E (IN XLRSL).
 /FTID-VAR_008220.
 G -> R (IN XLRSL).
 /FTID-VAR_008185.
 G -> W (IN XLRSL).
 /FTID-VAR_008221.
 C -> Y (IN XLRSL).
 /FTID-VAR_008222.
 W -> C (IN XLRSL).
 /FTID-VAR_008223.
 L -> F (IN XLRSL).
 /FTID-VAR_008224.
 L -> P (IN XLRSL).
 /FTID-VAR_008225.
 G -> V (IN XLRSL).
 /FTID-VAR_008226.
 I -> T (IN XLRSL).
 /FTID-VAR_008227.
 T -> A (IN XLRSL).
 /FTID-VAR_008228.
 G -> E (IN XLRSL).
 /FTID-VAR_008229.
 G -> R (IN XLRSL).
 /FTID-VAR_008230.
 R -> C (IN XLRSL).
 /FTID-VAR_008231.
 R -> G (IN XLRSL).
 /FTID-VAR_008232.
 R -> H (IN XLRSL).
 /FTID-VAR_008233.
 C -> W (IN XLRSL).
 /FTID-VAR_008234.
 D -> V (IN XLRSL).
 /FTID-VAR_008235.
 E -> D (IN XLRSL).
 /FTID-VAR_008236.
 E -> K (IN XLRSL).
 /FTID-VAR_008237.
 Y -> C (IN XLRSL).
 /FTID-VAR_008238.

Query Match	Best Local Match	Similarity	Conservative	Score	DB	Length	Indels	Gaps
223 YPOVCHRRIC-----TLRMELIGELSGC--SRLPMKSGHIDQVQTASSIFRLTM	272	11.88	33.28	267	1	224		
Db	34 YORACKDQGGPAPNALMSAGATSLD-IPCCYKHPKPGFGECEVTDPDQITCS-----NP	86						
Oy	273 DMF-----TWERRKARLDKQGRVYATSGHNSQSGWLQVLDLPVTVSIIYOGAKOEGH	327						
Db	87 EOYVGVSWTANKARKLNSQGFSCAMLSKFQDSQWLQIDIKETIKYSIIITOGCDD-145	145						
Oy	328 VOVSYSKYKAAVSNQGE-HMTVYODEKQRKDYFQSGNEQDNQTHRKRVIDPPIYARHRIPL	366						
Db	146 -EEMTGFVSQVYHTDEBLRMVITYKDY-TGNRRYFQNSDRISYQNLRLPISRIIRLP	203						
Oy	387 WMSWGRITLASRLGCTEE	405						
Db	204 LGHHVRIALRMELLECYSK	222						
RESULT 5								
DLK_HUMAN	STANDARD;	PRT;	383 AA.					
AC	P60370;							
DT	01-NOV-1995 (Rel. 32, Created)							
DT	01-FEB-1996 (Rel. 33, Last sequence update)							
DT	01-OCT-2000 (Rel. 40, Last annotation update)							
DE	DELTA-LIKE PROTEIN PRECURSOR (DLK) [CONTAINS: FETAL ANTIGEN 1 (FA1)].							
GN	DLK1 OR DLK.							
OS	Eukaryotes (Human).							
OC	Chordata; Chordata; Vertebrata; Euteleostomi;							
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.							
OC	NCBI_TaxID=9606;							
RP	SEQUENCE FROM N.A.							
RP	TISSUE-Adrenal gland;							
RX	MEDLINE=93179372; Pubmed=8095043;							
RA	Laborda J., Sauvillie E.A., Hoffman T., Notario V.;							
RT	"dlk, a putative mammalian homeotic gene differentially expressed in							
RT	small cell lung carcinoma and neuroendocrine tumor cell line.";							
RL	J. Biol. Chem. 268:3817-3820(1993).							
RP	SEQUENCE FROM N.A.							
RP	TISSUE-Placenta, and Adrenal gland;							
RX	MEDLINE=95226449; Pubmed=7711066;							
RA	Lee Y.L., Helman L., Hoffman T., Laborda J.;							
RT	"dlk, p62 and Pref-1 mRNAs encode similar proteins belonging to the							
RT	EGF-like superfamily. Identification of polymorphic variants of this							
RL	RNA.";							
RL	Biochim. Biophys. Acta 1261:223-232(1995).							
RP	SEQUENCE OF 24-383.							
RP	TISSUE-Amniotic fluid;							
RX	MEDLINE=95010145; Pubmed=7925474;							
RA	Jensen C.H., Krogsh T.N., Hoejrup P., Clausen P.P., Skjoedt K.;							
RA	Larsen L., T., Enghild J.J., Telser B.;							
RT	"Protein structure of fetal antigen 1 (FA1). A novel circulating							
RT	human epidermal growth-factor-like protein expressed in							
RT	neuroendocrine tumors and its relation to the gene products of dlk							
RT	and p62.";							

1	EMBL	212172	CAAT1613.1	-	EMBL	212172	CAAT1613.1	-
2	DR	EMBL	U13979	AAA75364.1	-	DR	EMBL	U13979
3	DR	EMBL	U13981	AAA75365.1	-	DR	EMBL	U13981
4	DR	HSSP	P00750	ITPG	-	DR	HSSP	P00750
5	DR	MM	176290	-	DR	MM	176290	-
6	DR	Interpro	IPR00152	-	DR	Interpro	IPR00152	-
7	DR	Pfam	PF00008	EGF_6	-	DR	Pfam	PF00008
8	DR	PROSITE	PS00010	ASX_HYDROXYL_1	-	DR	PROSITE	PS00010
9	DR	PROSITE	PS00022	EGF_1	5	DR	PROSITE	PS00022
10	KW	Antigen	Transmembrane	EGF_2	5	KW	Antigen	Transmembrane
11	KW	Alternative splicing	-	-	-	KW	Alternative splicing	-
12	FT	SIGNAL	1	23	FT	SIGNAL	1	23
13	FT	CHAIN	24	383	FT	CHAIN	24	383
14	FT	CHAIN	24	303	FT	CHAIN	24	303
15	FT	DOMAIN	24	303	FT	DOMAIN	24	303
16	FT	TRANSMEM	304	327	FT	TRANSMEM	304	327
17	FT	DOMAIN	328	383	FT	DOMAIN	328	383
18	FT	DOMAIN	24	55	FT	DOMAIN	24	55
19	FT	DOMAIN	53	86	FT	DOMAIN	53	86
20	FT	DOMAIN	88	125	FT	DOMAIN	88	125
21	FT	DOMAIN	127	168	FT	DOMAIN	127	168
22	FT	DOMAIN	170	206	FT	DOMAIN	170	206
23	FT	DOMAIN	208	245	FT	DOMAIN	208	245
24	FT	DISULFID	30	43	FT	DISULFID	30	43
25	FT	DISULFID	45	54	FT	DISULFID	45	54
26	FT	DISULFID	57	68	FT	DISULFID	57	68
27	FT	DISULFID	63	74	FT	DISULFID	63	74
28	FT	DISULFID	76	85	FT	DISULFID	76	85
29	FT	DISULFID	92	103	FT	DISULFID	92	103
30	FT	DISULFID	97	113	FT	DISULFID	97	113
31	FT	DISULFID	115	124	FT	DISULFID	115	124
32	FT	DISULFID	131	144	FT	DISULFID	131	144
33	FT	DISULFID	138	156	FT	DISULFID	138	156
34	FT	DISULFID	158	167	FT	DISULFID	158	167
35	FT	DISULFID	174	185	FT	DISULFID	174	185
36	FT	DISULFID	179	194	FT	DISULFID	179	194
37	FT	DISULFID	196	205	FT	DISULFID	196	205
38	FT	DISULFID	212	223	FT	DISULFID	212	223

```

FT DISULFID 217 233 BY SIMILARITY.
FT DISULFID 235 244 BY SIMILARITY.
FT CARBOHYD 94 94 O-LINKED.
FT CARBOHYD 100 100 N-LINKED (GLYCANC. . .).
FT CARBOHYD 143 143 O-LINKED.
FT CARBOHYD 163 163 O-LINKED (PARTIAL).
FT CARBOHYD 165 165 N-LINKED (GLYCANC. . .) (PARTIAL).
FT CARBOHYD 172 172 N-LINKED (GLYCANC. . .) (PARTIAL).
FT CARBOHYD 214 214 O-LINKED.
FT CARBOHYD 222 222 O-LINKED (PARTIAL).
FT CARBOHYD 251 251 O-LINKED (PARTIAL).
FT CARBOHYD 260 260 O-LINKED (PARTIAL).
FT CARBOHYD 229 301 MISSING (IN SHORT ISOFORM).
FT VARSPLIC 347 347 MISSING (IN CLONE HDLKAAG).
FT VARIANT 347 347 /FTID-VAR_002274.
FT CONFLICT 46 47 G -> HV (IN REF. 1).
FT CONFLICT 108 108 G -> D (IN REF. 3).
SQ SEQUENCE 383 AA; 41143 MW; 2C55C6E355B4415B CRC64;

Query Match 8.3%; Score 187; DB 1; Length 383;
Best Local Similarity 29.5%; Pred. No. 9.4e-08;
Matches 39; Conservative 16; Mismatches 41; Indels 36; Gaps 4;

QY 1 GPCPN--PCNNGCTEISEAYRGDTFIVGVCKCPRGNGHCOHNECEVEPKNGKI 58
DB 129 GPCVINGSPCGHGCTGVDEGRASHA-----SCLCPGSGNFCETIVANSCTPCENGCV 184
QY 59 CTDLVANVSCGCEGEMGRNC-----OYKC-----SGPLG 88
DB 189 CTDIGSGFRCRCAGFIDIKTSRPTVCASSPCQNGCTCLOHGVSEYELCPREFTLTC 244
QY 89 IEGGLISNOOT 100
DB 245 VKKRALSPQVY 256

RESULT 6
DIL_MOUSE STANDARD: PRT; 385 AA.
AC 009163; 007645; 062208;
DN 01-NOV-1995 (Rel. 32; Created)
DN 01-NOV-1995 (Rel. 32; Last sequence update)
DN 01-OCT-2000 (Rel. 40; Last annotation update)
DE DELTA-LIKE PROTEIN PRECURSOR (DLK) (PREDICTED FACTOR 1) (PREF-1)
DE (ADIPOCYTE DIFFERENTIATION INHIBITOR PROTEIN) (CONTAINS: FETAL ANTIGEN 1 (FAL)).
GN DLK1 OR DLK OR PREF1 OR SCP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS; TISSUE=Fibroblast;
RX Laborda J., Sausville E.A., Hoffman T., Notario V.,
RA Laborda J., Sausville E.A., Hoffman T., Notario V.,
RT "dlk, a putative mammalian homeotic gene differentially expressed in
RT small cell lung carcinoma and neuroendocrine tumor cell line.";
RL J. Biol. Chem. 268:3817-3820(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93272313; PubMed=8500166;
RA Smaas C.M., Sul H.S.;
RT "Pref-1, a protein containing EGF-like repeats, inhibits adipocyte
RT differentiation.";
RN [3]
RP SEQUENCE FROM N.A.; AND VARIANT LYS-347 DEL.
RX TISSUE=Adrenal gland, and Placenta;
RX MEDLINE=95226449; PubMed=7711066;
RA Lee T.L., Helman L., Hoffman T., Laborda J.;
RT "dlk, p62 and Pref-1 mRNAs encode similar proteins belonging to the

```

```

RT EGF-like superfamily. Identification of polymorphic variants of this
RT RNA.";
RL Biochim. Biophys. Acta 1261:223-232(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA Maruyama K., Nishijima S., Kurumitsu S., Ichikawa A., Masuda E.,
RA Takemoto T., Kodama H., Kawashima H.;
RL Submitted (Aug-1993) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-8 FROM N.A.; AND ALTERNATIVE SPLICING.
RX MEDLINE=94325292; PubMed=7519443;
RA Smaas C.M., Green D., Sul H.S.;
RT "Structural characterization and alternate splicing of the gene
RT encoding the preadipocyte EGF-like protein pref-1.";
RL Biochemistry 33:9257-9265(1994).
RN [6]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=97234561; PubMed=9118998;
RA Krogh T.N., Bachmann E., Teisner B., Skjodet R., Hoesrup P.;
RT "Glycosylation analysis and protein structure determination of murine
RT delta-like protein (dlp) - the circulating gene product of the
RT stromal-cell-derived protein 1 (SCP-1) CDMS.";
RL Eur. J. Biochem. 244:331-342(1997).
RN [7]
RP FUNCTION: MAY HAVE A ROLE IN NEUROENDOCRINE DIFFERENTIATION.
CC -1- INHIBITS ADIPOCYTE DIFFERENTIATION.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: A (SHOWN HERE), B, C,
CC C2, D AND D2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PIV: N AND D2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z12171; CAA8162.1; -;
DR EMBL; U15980; AAB60495.1; -;
DR EMBL; L12721; AAA37175.1; -;
DR EMBL; S71340; -; NOT_ANNOTATED_CDS.
DR HSSP; D16847; BAA04121.1; -;
DR HSSP; P00740; IIXA.
DR GlycoSiteDB; Q09163; -;
DR MGD; MGI:94900; Dlk1.
DR InterPro; IPR000152; -;
DR InterPro; IPR000561; -;
DR Pfam; PF00008; EGF; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 6.
KW Antigen; Transmembrane; EGF-like domain; Repeat; Glycoprotein; Signal;
KW Alternative splicing.
FT SIGNAL 1 23
FT CHAIN 24 385 BY SIMILARITY.
FT CHAIN 24 305 DELTA-LIKE PROTEIN.
FT CHAIN 24 305 FETAL ANTIGEN 1.
FT DOMAIN 306 329 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 330 385 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 24 55 EGF-LIKE 1.
FT DOMAIN 53 86 EGF-LIKE 2.
FT DOMAIN 88 125 EGF-LIKE 3.
FT DOMAIN 127 168 EGF-LIKE 4.
FT DOMAIN 172 208 EGF-LIKE 5.
FT DOMAIN 210 247 EGF-LIKE 6.
FT DISULFID 26 37 BY SIMILARITY.
FT DISULFID 30 43 BY SIMILARITY.
FT DISULFID 45 54 BY SIMILARITY.
FT DISULFID 57 68 BY SIMILARITY.

```

```

FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 76 85 BY SIMILARITY.
FT DISULFID 92 103 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 156 BY SIMILARITY.
FT DISULFID 158 167 BY SIMILARITY.
FT DISULFID 176 187 BY SIMILARITY.
FT DISULFID 181 196 BY SIMILARITY.
FT DISULFID 198 207 BY SIMILARITY.
FT DISULFID 214 225 BY SIMILARITY.
FT DISULFID 219 235 BY SIMILARITY.
FT DISULFID 237 246 BY SIMILARITY.
FT CARBOHYD 94 94 O-LINKED.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT CARBOHYD 216 216 O-LINKED.
FT CARBOHYD 224 224 O-LINKED.
FT CARBOHYD 258 258 O-LINKED.
FT CARBOHYD 267 267 O-LINKED (PARTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 MISSING (IN ISOFORM D).
FT VASAPLIC 211 303 MISSING (IN ISOFORM D2).
FT VASAPLIC 211 303 MISSING (IN ISOFORM B).
FT VASAPLIC 231 303 MISSING (IN ISOFORM C).
FT VASAPLIC 231 303 MISSING (IN ISOFORM C2).
FT VARIANT 347 347 MISSING.
FT CONFLICT 250 250 R -> P (IN REF. 2).
FT CONFLICT 320 385 VLGVAVLFLNCEIWEVNLRYNTPKRNKLLLOYSCEE
LAVNIIEPKIMDETEAGDEEI -> CAMPMWSSSTS
AKGPCPTCATTTTCFARRRSCCTTARSRSISSPRRLT
(IN REF. 2).
TF -> ML (IN REF. 4).
SQ CONFLICT 344 345 E79864FEAFAFAFI CRC64;

```

```

Query Match
Best Local Similarity 33.7%; Score 179; DB 1; Length 385;
Matches 36; Conservative 14; Mismatches 46; Indels 12; Gaps 4;
QY 1 GPCPN--PCNGGCEISEAHRDPTFGVCKCPGFGNHCQ--NNINECEVERCKNG 56
DB 129 GPCVINSFGGACVDEGASHA-----SCLCPEFGFSGNFCVVAATNCTPFCEND 184
QY 57 GICDVLVANSCEGCEFGFNGRCYK-----CSEGLDIBEGITISNOQT 100
DB 185 GVCIDIGDRCRCRPAFGVDTKTSRPVSNCSAGSCONGSTCLQHTQVS 232

```

```

RESULT 7
WIFI_MOUSE STANDARD: PRT; 379 AA.
ID WIFI_MOUSE
AC 09WDAL;
DT 01-OCT-2000 (Rel. 40; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE WNT INHIBITORY FACTOR 1 PRECURSOR (WIF-1).
GN WIFI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RA MEDLINE=99215557; PubMed=10201374;
RA Hsieh J.-C., Kodjabachian L., Rebber M.L., Ratner A.,
RA Smallwood P.M., Samos C.H., Nusse R., David I.B., Nathans J.;
RA "A new secreted protein that binds to Wnt proteins and inhibits their
RT activities";
RT Nature 398:431-436(1999).

```

```

CC -1- FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY
CC BE INVOLVED IN MESODERM SEGMENTATION.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: EXPRESSION HIGHEST IN HEART AND LUNG. LOWER IN
CC BRAIN AND EYE.
CC -1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF122923; AAD25403.1; -.
CC HSSP: P00740; 11XA.
CC MGD: MG1:1344332; WIF1.
CC InterPro: IPR000561; -.
CC InterPro: IPR002049; -.
CC Pfam: PF00008; EGF: 5.
CC PRINTS: PR00011; EGF-LAMININ.
CC PROSITE: PS00022; EGF_1; 5.
CC PROSITE: PS01186; EGF_2; 4.
CC Repeat: EGF-like domain; Signal: Developmental protein.
CC
CC SIGNL 1 28
CC CHAIN 29 379 WNT INHIBITORY FACTOR 1.
CC DOMAIN 177 208 EGF-LIKE 1.
CC DOMAIN 209 240 EGF-LIKE 2.
CC DOMAIN 241 272 EGF-LIKE 3.
CC DOMAIN 273 304 EGF-LIKE 4.
CC DOMAIN 305 336 EGF-LIKE 5.
CC DISULFID 177 186 POTENTIAL.
CC DISULFID 182 192 POTENTIAL.
CC DISULFID 198 200 POTENTIAL.
CC DISULFID 209 218 POTENTIAL.
CC DISULFID 214 224 POTENTIAL.
CC DISULFID 230 232 POTENTIAL.
CC DISULFID 241 250 POTENTIAL.
CC DISULFID 246 256 POTENTIAL.
CC DISULFID 262 264 POTENTIAL.
CC DISULFID 273 282 POTENTIAL.
CC DISULFID 278 288 POTENTIAL.
CC DISULFID 294 296 POTENTIAL.
CC DISULFID 305 314 POTENTIAL.
CC DISULFID 310 320 POTENTIAL.
CC DISULFID 326 328 POTENTIAL.
CC CARBOHYD 88 88 POTENTIAL.
CC CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 379 AA; 41590 MM; E3765F2642B2BC9A CRC64;

```

```

Query Match
Best Local Similarity 6.2%; Score 141.5; DB 1; Length 379;
Matches 35; Conservative 6; Mismatches 30; Indels 33; Gaps 5;
QY 5 PNPCHNGGCEISEAYRNDPTFGVCKCPGFGNHCQNNINECEVERCKNGICTDLVA 64
DB 183 PGCCRRGPGGNCNRR-----VCECPDGFYGPCHCKAL--C-IPRCWNGGLC---VT 226
QY 65 NYSCECPGCEFGFNGRCYK-----KQSGPLDIEG 91
DB 227 PGFCICPPGFGVNCMDKNCSTCFNGICTFYGRKICPPGLEG 270
RESULT 8
DIS2_DICDI STANDARD: PRT; 257 AA.
ID DIS2_DICDI
AC P42530;
DT 01-NOV-1995 (Rel. 32; Created)
DT 01-NOV-1995 (Rel. 32; Last sequence update)
DT 15-JUL-1998 (Rel. 36; Last annotation update)
DE DISCORDIN II.

```

```

OS Homo sapiens (Human) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TextID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99215557; PubMed=10201374;
RA Heish T.-C., Kodjabachian L., Robert M.L., Rattner A.,
RA Smallwood P.M., Samos C.H., Nussle R., David I.B., Nathans J.;
RT "A new secreted protein that binds to Wnt proteins and inhibits their
RT activities."
RL Nature 399:431-436(1999).
CC -!- FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY
CC BE INVOLVED IN MESODERM SEGMENTATION.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC -!-
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF122922; AAD25402.1; -.
DR MIM: 605186; -.
DR HSSP: P00743; 1CCF.
DR InterPro: IPR000561; -.
DR InterPro: IPR002049; -.
DR Pfam: PF00008; EGF: 5.
DR PRINTS: PR00011; EGF-LAMININ.
DR PROSITE: PS00022; EGF_1: 5.
DR PROSITE: PS01186; EGF_2: 4.
KV Repeat: EGF-like domain; Signal: Developmental protein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 379 WNT INHIBITORY FACTOR 1.
FT DOMAIN 177 208 EGF-LIKE 1.
FT DOMAIN 209 240 EGF-LIKE 2.
FT DOMAIN 241 272 EGF-LIKE 3.
FT DOMAIN 273 304 EGF-LIKE 4.
FT DOMAIN 305 336 EGF-LIKE 5.
FT DISULFID 177 186 POTENTIAL.
FT DISULFID 182 192 POTENTIAL.
FT DISULFID 198 200 POTENTIAL.
FT DISULFID 209 218 POTENTIAL.
FT DISULFID 214 224 POTENTIAL.
FT DISULFID 230 232 POTENTIAL.
FT DISULFID 241 250 POTENTIAL.
FT DISULFID 246 256 POTENTIAL.
FT DISULFID 262 264 POTENTIAL.
FT DISULFID 273 282 POTENTIAL.
FT DISULFID 278 288 POTENTIAL.
FT DISULFID 294 296 POTENTIAL.
FT DISULFID 305 314 POTENTIAL.
FT DISULFID 310 320 POTENTIAL.
FT DISULFID 326 328 POTENTIAL.
FT CARBOHD 88 88 POTENTIAL.
FT CARBOHD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 379 AA; 41512 MW; 27782370A266E784 CRC64;

Query Match 6.2%; Score 140.5; DB 1; Length 379;
Best Local Similarity 32.7%; Pred. No. 0.00045;
Matches 34; Conservative 7; Mismatches 30; Indels 33; Gaps 5;

OY 5 PNPCHNGTCEISEAYRGDFIGYVCKPRGNGINCHNINCEVEBCKNGSICITDLYA 64
DB 183 PGKNGCNCFENRR-----ICECDGPHPEKAL--C-TRCKNNSGLC--VT 226
OY 65 NYSCECPGEFMRNCY-----KCSGPLGEG 91
DB 227 PGFCICPGPYGVNCDKANGSTTCFNGSGTCEYPCKCIQCPGLG 270

```

```

RESULT 10
ID DISC_DICDI STANDARD: PRT: 149 AA.
AC P02887:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DISCOIDIN I, D CHAIN (FRAGMENT).
GN DSCD.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82170475; PubMed=6279874;
RA Poole S., Firtel R.A., Lamar E., Rowekamp W.;
RT "Sequence and expression of the discoidin I gene family in
RT Dictyostelium discoideum."
RL J. Mol. Biol. 153:273-289(1981).
RN [2]
RP CELL ATTACHMENT SITE.
RX MEDLINE=85074463; PubMed=6509552;
RA Springer W.R., Cooper D.N.W., Battondes S.H.;
RT "Discoidin I is implicated in cell-substratum attachment and ordered
RT cell migration of Dictyostelium discoideum and resembles
RT fibronectin."
RL Cell 39:557-564(1984).
CC -1- FUNCTION: GALACTOSE- AND N-ACETYLGALACTOSAMINE-BINDING LECTIN. MAY
CC PLAY A ROLE IN CELL-SUBSTRATUM ADHESION RATHER THAN IN CELL-CELL
CC ADHESION. MAY BE NECESSARY FOR THE MAINTENANCE OF NORMAL ELONGATE
CC MORPHOLOGY DURING AGGREGATION.
CC -1- SUBUNIT: TETRAMER OF FOUR DIFFERENT CHAINS (A TO D).
CC -1- TISSUE SPECIFICITY: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: STALK CELLS.
CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01285; AAA33200.1; -
DR PIR: A03383; DLDOLD.
DR DictyDB: DD02002; dscd.
DR InterPro: IPR000421; -
DR Pfam: PF00754; F5_F8_Type_C; 1.
DR PROSITE: PS01285; FA58C_1; 1.
DR PROSITE: PS01286; FA58C_2; FALSE_NEG.
KW Lectin; Cell adhesion.
FT SITE 79 81 CELL ATTACHMENT SITE.
FT NON_TER 149 149
SO SEQUENCE 149 AA; 16746 MW; 1B8E0F20B29E9A1 CRC64;

Query Match 5.9%; Score 134.5; DB 1; Length 149;
Best Local Similarity 34.6%; Pred. No. 0.00046;
Matches 37; Conservative 11; Mismatches 52; Indels 7; Gaps 2;

```

```

ID DISC_DICDI STANDARD: PRT: 253 AA.
AC P02887:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DISCOIDIN I, C CHAIN AND B CHAIN.
GN DSCB AND DSCC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A. (C CHAIN AND B CHAIN 1-149).
RX MEDLINE=82170475; PubMed=6279874;
RA Poole S., Firtel R.A., Lamar E., Rowekamp W.;
RT "Sequence and expression of the discoidin I gene family in
RT Dictyostelium discoideum."
RL J. Mol. Biol. 153:273-289(1981).
RN [2]
RP CELL ATTACHMENT SITE.
RX MEDLINE=85074463; PubMed=6509552;
RA Springer W.R., Cooper D.N.W., Battondes S.H.;
RT "Discoidin I is implicated in cell-substratum attachment and ordered
RT cell migration of Dictyostelium discoideum and resembles
RT fibronectin."
RL Cell 39:557-564(1984).
CC -1- FUNCTION: GALACTOSE- AND N-ACETYLGALACTOSAMINE-BINDING LECTIN. MAY
CC PLAY A ROLE IN CELL-SUBSTRATUM ADHESION RATHER THAN IN CELL-CELL
CC ADHESION. MAY BE NECESSARY FOR THE MAINTENANCE OF NORMAL ELONGATE
CC MORPHOLOGY DURING AGGREGATION.
CC -1- SUBUNIT: TETRAMER OF FOUR DIFFERENT CHAINS (A TO D).
CC -1- TISSUE SPECIFICITY: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: STALK CELLS.
CC -1- SIMILARITY: THE SEQUENCE OF RESIDUES 1-149 OF THE B CHAIN IS
CC IDENTICAL WITH THAT OF THE C CHAIN.
CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01284; AAA33199.1; -
DR PIR: A03382; DLDOLD.
DR DictyDB: DD02001; dsc.
DR InterPro: IPR000421; -
DR Pfam: PF00754; F5_F8_Type_C; 1.
DR PROSITE: PS01285; FA58C_1; 1.
DR PROSITE: PS01286; FA58C_2; FALSE_NEG.
KW Lectin; Cell adhesion.
FT SITE 79 81 F5/8 TYPE C.
FT NON_TER 149 149 CELL ATTACHMENT SITE.
SO SEQUENCE 253 AA; 26392 MW; 875EA94811F55444 CRC64;

Query Match 5.9%; Score 134.5; DB 1; Length 253;
Best Local Similarity 34.6%; Pred. No. 0.00085;
Matches 37; Conservative 11; Mismatches 52; Indels 7; Gaps 2;

```

```

RESULT 12
PRZ_BOVIN STANDARD; PRT: 396 AA.
AC P00744;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE VITAMIN K-DEPENDENT PROTEIN Z.
GN PROZ.
OS Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN 1;
RP SEQUENCE.
RX MEDLINE=85204370; PubMed=3888670;
RA Hoelrup P., Jensen M.S., Petersen T.E.;
RT "Amino acid sequence of bovine protein Z: a vitamin K-dependent
RT serine protease homolog."
RL FEBS Lett. 184:333-338(1985).
RN 12;
RP STRUCTURE OF CARBOHYDRATE ON SER-53.
RX MEDLINE=90062160; PubMed=2511201;
RA Nishimura H., Kawabata S., Kistiel W., Hase S., Ikenaka T., Takao T.,
RA Shimonishi Y., Iwanaga S.;
RT "Identification of a disaccharide (Xyl-Glc) and a trisaccharide
RT (Xyl2-Glc) O-glycosidically linked to a serine residue in the first
RT epidermal growth factor-like domain of human factors VII and IX and
RT protein Z and bovine protein Z."
RL J. Biol. Chem. 264:20320-20325(1989).
RN 13;
RP STRUCTURE OF CARBOHYDRATE ON SER-53.
RX MEDLINE=91344709; PubMed=2129367;
RA Iwanaga S., Nishimura H., Kawabata S., Kistiel W., Hase S., Ikenaka T.;
RT "A new trisaccharide sugar chain linked to a serine residue in the
RT first EGF-like domain of clotting factors VII and IX and protein Z."
RL Adv. Exp. Med. Biol. 281:121-131(1990).
CC 1- FUNCTION: SINGLE-CHAIN PLASMA GLYCOPROTEIN OF UNKNOWN FUNCTION.
CC 1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC 1- TISSUE SPECIFICITY: PLASMA.
CC 1- SIMILARITY: ALTHOUGH HOMOLOGOUS WITH THE VITAMIN K-DEPENDENT
CC CLOTTING FACTORS, IT HAS LOST TWO OF THE ESSENTIAL CATALYTIC
CC RESIDUES AND HAS NO ENZYMIC ACTIVITY.
CC 1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC PIR: A2171; KKB02.
DR HSSP: P00740; ICER.
DR MEROPS: S01.979; -.
DR GlycoSuiteDB: P00744; -.
DR InterPro: IPR000152; -.
DR InterPro: IPR000294; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR001254; -.
DR InterPro: IPR002383; -.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00594; Glu_1.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00001; GLABLOO.
DR PROSITE: PS00010; ASX_HYDROXYL_1.
DR PROSITE: PS00011; GLU CARBOXYLATION_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01166; EGF_2; 1.
KW Plasma; Glycoprotein; Gamma-carboxyglutamic acid; Hydroxylation;
KW Calcium; Serine protease homolog; Vitamin K; EGF-like domain.
FT DOMAIN 47 83 EGF-Like 1.
FT DOMAIN 85 126 EGF-Like 2.
FT DOMAIN 143 396 TRYPSIN-LIKE.
FT MOD_RES 7 7 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 8 8 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 11 11 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 15 15 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 17 17 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 20 20 GAMMA-CARBOXYGLUTAMIC ACID.

```

```

FT MOD_RES 21 21 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 26 26 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 27 27 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 30 30 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 33 33 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 36 36 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 40 40 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 64 64 HYDROXYLATION.
FT DISULFID 51 62 BY SIMILARITY.
FT DISULFID 56 71 BY SIMILARITY.
FT DISULFID 73 82 BY SIMILARITY.
FT DISULFID 89 101 BY SIMILARITY.
FT DISULFID 97 110 BY SIMILARITY.
FT DISULFID 112 125 BY SIMILARITY.
FT CARBOHYD 53 53 O-LINKED (GLC. . .).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 396 AA; 43112 MW; 04C5D7A35849B116 CRC64;

Query Match 5.48; Score 122.5; DB 1; Length 396;
Best Local Similarity 31.34; Pred. No. 0.013;
Matches 31; Conservative 16; Mismatches 33; Indels 19; Gaps 7;

OY 2 PCPNNCHNGGTCETSEAYRNDPFGYCKCPGRFGHGHCHNINEC--EVEPKNGCI 58
DB 50 PCASQCLNNGSCQ-----DSIRGACTCAPGYEGNCAPASECHPLRLDGCQH--F 100
OY 59 CTDLVANSCECP--GEPMG--QKCS--GLGIE 90
DB 101 CYPEPSYTCARSHKRGDRNSCLPHRCACGLGFE 139

RESULT 13
DISA.DICDI STANDARD; PRT: 253 AA.
AC P02886;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DISCOLDIN I, A CHAIN.
GN DSCA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE=82170475; PubMed=6279874;
RA Poole S., Firtel R.A., Lamar E., Rowekamp W.;
RT "Sequence and expression of the discoidin I gene family in
RT Dictyostelium discoideum."
RL J. Mol. Biol. 153:273-289(1981).
RN 12;
RP REVISIONS, SEQUENCE FROM N.A.
RX MEDLINE=83059644; PubMed=6754951;
RA Jellingshaus U., Schaeferle U., Schmid W., Rowekamp W.;
RT "Transcription of a dictyostelium discoidin-I gene in yeast
RT alternative promoter sites used in two different eukaryotic cells."
RL J. Mol. Biol. 153:623-636(1982).
RN 13;
RP SEQUENCE OF 1-40 FROM N.A.
RX MEDLINE=82235698; PubMed=6284373;
RA Devine J.M., Tsang A.S., Williams J.G.;
RT "Differential expression of the members of the discoidin I multigene
RT family during growth and development of Dictyostelium discoideum."
RL Cell 28:793-800(1982).
RN 14;
RP CELL ATTACHMENT SITE.
RX MEDLINE=85074463; PubMed=6509552;
RA Springer W.R., Cooper D.N.W., Barondes S.H.;
RT "Discoidin I is implicated in cell-substratum attachment and ordered

```

[illegible]

```

FT MOD_RES 80 80 GAMMA-CARBOXYGLUTAMIC ACID
FT DISULFID 104 104 HYDROXYLATION (BY SIMILARITY).
FT DISULFID 91 102 BY SIMILARITY.
FT DISULFID 96 111 BY SIMILARITY.
FT DISULFID 113 122 BY SIMILARITY.
FT DISULFID 129 141 BY SIMILARITY.
FT DISULFID 137 150 BY SIMILARITY.
FT DISULFID 152 165 BY SIMILARITY.
FT CARBOHYD 93 93 O-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 400 AA: 44744 MW: 78BD2DCC48860268 CRC64:

Query Match 5.28: Score 117; DB 1; Length 400;
Best Local Similarity 31.98; Pred. No. 0.035; Indels 12; Gaps 3;
Matches 23; Conservative 11; Mismatches 26;

OY 2 CTPNPNCHNGTCETSEAYRQDFITGYVCKCPGFGNGHICOHNIKECEVE--PCKNGGI 58
DB 90 PCISQPCPLHNGSCQ-----DSIMGYTCCTSPGYEGSGNCELANKECHPERTDCQH--F 140
OY 59 CTDIVANYSCEC 70
DB 141 CLPQGESEYTCSC 152

RESULT 15
ID FA10_TROCA STANDARD: PRT: 376 AA.
AC P81428:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DS CONGULATION FACTOR X (EC 3.4.21.6) (TROCKARIN PROTHROMBIN ACTIVATOR).
OS Tropidochelis carinata (Australian rough-scaled snake).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Tropidochelis.
OX NCBI_TaxID=100989;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RX TISSUE-Venom;
RA MEDLINE=99326314; PubMed=10397729;
RA Joseph J.S., Chung M.C.M., Jayaseelan K., Kini R.N.;
RT "Amino acid sequence of trocacin, a prothrombin activator from
RT Tropidochelis carinata venom: its structural similarity to coagulation
RT factor Xa."
RL Blood 94:621-631(1999).
CC -1- FUNCTION: FACTOR XA IS A VITAMIN K-DEPENDENT GLYCOPROTEIN THAT
CC CONVERTS PROTHROMBIN TO THROMBIN IN THE PRESENCE OF FACTOR VA,
CC CALCIUM, AND PHOSPHOLIPID DURING BLOOD CLOTTING. ACTS AS A TOXIN
CC IN VENOM.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-1-THR AND THEN
CC ARG-1-ILE BONDS IN PROTHROMBIN TO FORM THROMBIN.
CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR
CC AND ARE HELD TOGETHER BY 1 OR MORE DISULFIDE BONDS.
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMAIC CARBOXYLATION OF SOME
CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC CALCIUM.
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR MEROPS, S01.216;
DR InterPro, IPR000152;
DR InterPro, IPR000294;
DR InterPro, IPR000561;
DR InterPro, IPR001254;
DR InterPro, IPR001314;
DR InterPro, IPR001438;
DR InterPro, IPR001881;
DR InterPro, IPR002383;
DR PRINTS, PR00001; GLABLOOD.
DR PRINTS, PR00010; EGFBLD.
DR PRINTS, PR00722; CHYMOTRYPSIN.
DR PROSITE, PS00010; ASX HYDROXYL. 1.
DR PROSITE, PS00011; GLU CARBOXYLATION; 1.
DR PROSITE, PS00022; EGF 1; 1.
DR PROSITE, PS00187; EGF CA; 1.
DR PROSITE, PS00134; TRYPSIN_HIS; 1.
DR PROSITE, PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW Zymogen; EGF-like domain; Repeat.
FT CHAIN 1 141
FT MOD_CONS 141 142
FT CHAIN 142 376
FT CHAIN 142 376
FT DOMAIN 50 81
FT DOMAIN 89 124
FT DOMAIN 142 376
FT MOD_RES 6 6
FT MOD_RES 7 7
FT MOD_RES 14 14
FT MOD_RES 14 14
FT MOD_RES 16 16
FT MOD_RES 19 19
FT MOD_RES 20 20
FT MOD_RES 25 25
FT MOD_RES 26 26
FT MOD_RES 29 29
FT MOD_RES 32 32
FT MOD_RES 32 32
FT MOD_RES 33 35
FT MOD_RES 63 63
FT ACT_SITE 183 183
FT ACT_SITE 229 229
FT ACT_SITE 326 326
FT ACT_SITE 326 326
FT DISULFID 17 22
FT DISULFID 50 61
FT DISULFID 72 81
FT DISULFID 89 100
FT DISULFID 89 109
FT DISULFID 96 124
FT DISULFID 111 124
FT DISULFID 132 249
FT DISULFID 148 153
FT DISULFID 168 184
FT DISULFID 297 311
FT DISULFID 322 350
FT CARBOHYD 52 52
FT CARBOHYD 186 186
SQ SEQUENCE 376 AA: 42455 MW: F5A5C7F89BA561C1 CRC64:

Query Match 4.48: Score 99.5; DB 1; Length 376;
Best Local Similarity 21.5%; Pred. No. 0.79; Indels 61; Gaps 12;
Matches 45; Conservative 25; Mismatches 78;

OY 3 CTPNPNCHNGTCETSEAYRQDFITGYVCKCPGFGNGHICOHNIKECEVEPCCKNG--I 58
DB 50 CSSNCHVGRGCK-----DGISGYTCCTCLPNYEGKCKEYVQSCRYD--NONGCHWF 99
OY 59 CTDIVANYSCEGGERM-----GRNC-----QYKSGPLGIGGIIISQDTASSYHRLAFG 110
DB 100 CKRQSEIQCSCAESRILGVDSHCVAGSDSCORNIKANIKIVNMOCKLGECE----- 153
OY 111 LQKPTPYARL-NKG-----LINMTAENDRKRMKPTQITLQKKRMYGVITQGS 159
DB 154 -----PQDAVLINKEGVEPCGCTIISPIHVLTAH-----CLMQTKSVKETHRLSV 200

```


Wed May 23 08:50:22 2001

us-09-237-981-14_copy_108_513.clo.rsp

Page 13

OY * 160 AK-----RIGSPEYI---KEYKIASND 179
| : | : : |
Db 201 DKIVYHTKFPNPNYYVHONFDRVADYD 229

Search completed: May 23, 2001, 08:37:41
Job time: 122 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:35:19 ; Search time 27.1 Seconds
(without alignments)
1755.956 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513
Perfect score: 2264
Sequence: 1 GPCFPNCHNGTCEISEAV.....MSWGRITLASELLGCTERE 406

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 284491

Minimum DB seq length: 0
Maximum DB seq length: 406

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPRMBL.15:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mbc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.todent:*
12: sp.unclassified:*
13: sp.vertibrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	933.5	41.2	363	6	077718	Q17718 equus caball
2	464.5	20.5	216	4	014286	Q14286 mus musculu
3	428.5	18.9	192	11	Q9QW01	Q9QW01 mus musculu
4	279	12.3	224	11	Q9RLM6	Q9RLM6 mus musculu
5	224.5	9.9	156	5	Q26661	Q26661 stronglylloe
6	220	9.7	406	5	Q25059	Q25059 heliocidari
7	202.5	9.1	78	4	P78328	P78328 homo sapien
8	199.5	8.9	263	4	Q99734	Q99734 homo sapien
9	197.5	8.7	387	11	Q99740	Q99740 mus musculu
10	189.5	8.4	308	6	Q46370	Q46370 bos taurus
11	185	8.2	383	11	Q70534	Q70534 rattus norv
12	185	8.2	383	11	Q62779	Q62779 rattus norv
13	182.5	8.1	192	4	Q9UDM2	Q9UDM2 homo sapien
14	182.5	8.1	315	4	Q9UPK6	Q9UPK6 homo sapien
15	176.5	7.8	252	4	Q9NTE1	Q9NTE1 homo sapien
16	164	7.2	305	6	Q9N028	Q9N028 macaca fasc
17	164	6.5	355	4	Q9Y506	Q9Y506 homo sapien
18	148	6.4	323	5	Q9YMG4	Q9YMG4 drosophila
19	146					

20	144	6.4	247	5	Q9VME8	Q9VME8 drosophila
21	141.5	6.2	379	11	Q9W0A1	Q9W0A1 mus musculu
22	141	6.2	349	5	Q17377	Q17377 caenorhabd
23	140.5	6.2	379	4	Q9Y5W5	Q9Y5W5 homo sapien
24	139.5	5.9	78	4	Q9U119	Q9U119 homo sapien
25	134.5	5.9	364	4	Q14089	Q14089 homo sapien
26	134	5.9	403	5	Q18375	Q18375 drosophila
27	133	5.9	374	13	Q9W6F8	Q9W6F8 xenopus lae
28	133	5.9	379	5	Q19922	Q19922 caenorhabd
29	127.5	5.6	220	11	Q63404	Q63404 rattus norv
30	124	5.5	51	4	Q9U005	Q9U005 homo sapien
31	124	5.5	378	13	Q9W6F9	Q9W6F9 brachydanio
32	119.5	5.3	293	4	Q99944	Q99944 homo sapien
33	118.5	5.2	144	13	Q12973	Q12973 gallus gall
34	118	5.2	176	13	Q9PUC7	Q9PUC7 gallus gall
35	114.5	5.1	273	4	Q9UHF1	Q9UHF1 homo sapien
36	113.5	5.0	259	5	Q93519	Q93519 caenorhabd
37	113.5	5.0	293	11	Q35447	Q35447 mus musculu
38	111	4.9	399	4	Q9N015	Q9N015 homo sapien
39	110.5	4.9	135	13	Q9PUC8	Q9PUC8 gallus gall
40	110.5	4.8	379	11	Q35883	Q35883 rattus norv
41	108.5	4.8	297	11	Q08743	Q08743 mus musculu
42	108.5	4.8	372	5	Q21756	Q21756 caenorhabd
43	107.5	4.7	305	1	Q9U2M0	Q9U2M0 pyrococcus
44	106.5	4.7	106	4	Q9UN93	Q9UN93 homo sapien
45	106.5	4.7	308	4	Q9UN94	Q9UN94 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	363 AA.
ID	077718			
AC	077718			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	SPERM-MEMBRANE ASSOCIATED PROTEIN P47 (FRAGMENT).			
OS	Equus caballus (Horse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.			
RN	NCBI_TaxID=9796;			
RP	SEQUENCE FROM N.A.			
RA	TISSUE=TESTIS;			
RA	Genzel M., Roepfer-Petersen E.;			
RL	Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AJ010121; CAA09010.1;			
DR	INTERPRO: IPR000421;			
DR	INTERPRO: IPR000561;			
DR	PEAW: PF00008; EGF_1;			
DR	PEAW: PF00754; F5_E8_Type_C; 2.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.			
DR	PROSITE: PS01186; EGF_2; 2.			
DR	PROSITE: PS01285; FA58C_1; 1.			
FT	Sperm.			
FT	NON_TER			
FT	NON_TER			
SO	SEQUENCE			
	363 AA:	40744 MW:	1FBB6395AF32338D CRC64;	
Query Match	41.2%;	Score 933.5;	DB 6:	Length 363;
Best Local Similarity	44.2%;	Pred. No. 9.1e-76;		
Matches 175;	Conservative 63;	Mismatches 95;	Indels 63;	Gaps 5;
QY	1 GPCFPNCHNGTCEISEAVRGDPFVGYVCKCPKPFNGIHQNHINECEVEPKNGKIGIC 59			
DB				
DB	30 GPCFPNCHNGTCEISEAVRGDPFVGYVCKCPKPFNGIHQNHINECEVEPKNGKIGIC 74			
QY	60 TDIVANVSCCEPDEFGRNCOYKSGPLIGEGGIIISNOQITASSTHRALGLOKWPYYA 119			
DB				
DB	75 -----CAMPICMETGAIADAAISASVYGFGLRWVPEIA 111			


```

DB 145 VNMVTCDCLAGFTGSCNFTNINECASNPCLNGS 177
RESULT 7
ID P78328 PRELIMINARY; PRT; 78 AA.
AC P78328:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HPA7 PROTEIN (FRAGMENT).
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Esslin M.;
RL Thesis (1996). University of Hannover, Hannover, Germany.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA MEDLINE=98206817; PubMed=9546740;
RA Esslin M., Vogel T., Calvete J.J., Thole H.H., Schmidtke J.,
RA Matsuda T., Roepfer-Petersen E.;
RT "Molecular cloning and characterization of P47, a novel boar sperm-
RT associated zona pellucida-binding protein homologous to a family of
RT mammalian secretory proteins."
RL Biol. Reprod. 58:1057-1064(1998).
DR EMBL: Y11718; CAW2405.1; -
DR INTERPRO: IPR000421; -
DR PROSITE: PS01286; PAS8C_2. 1.
KW Sperm
FT NON_TER 1
FT NON_TER 78
FT SEQUENCE 78 AA; 8696 MW; 7EBBD64D59465DDA CRC64;
Query Match 9.18; Score 205; DB 4; Length 78;
Best Local Similarity 54.38; Pred. No. 2.7e-11;
Matches 38; Conservative 11; Mismatches 21; Indels 0; Gaps 0;
QY 187 KYGTVEDNVEFGNIDNFTFANSPTPIKAQYVRLPYQVCRHCTLRMLLGCCLSGCS 246
DB 6 QVAGRSQDRIPTGNVNSGLKINLFTPLETYVRLPEITICRQCTLRFLCCLSGCT 65
QY 247 EPIGKRSCHI 256
DB 66 EPIGLKDNMTI 75
RESULT 8
ID Q99734 PRELIMINARY; PRT; 263 AA.
AC Q99734:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE NOTCH2 TRANSMEMBRANE PROTEIN (FRAGMENT).
GN NOTCH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T CELL;
RA Lemasson I., Devaux C., Mesnard J.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U77493; ARI9224.1; -
DR HSSP: P00736; IAFQ.

```

```

DR INTERPRO: IPR000083; -
DR INTERPRO: IPR000152; -
DR INTERPRO: IPR000561; -
DR INTERPRO: IPR000742; -
DR INTERPRO: IPR001010; -
DR INTERPRO: IPR001389; -
DR INTERPRO: IPR001881; -
DR INTERPRO: IPR001947; -
DR INTERPRO: IPR002049; -
DR PAM: PF00008; EGF_6.
DR PRINTS: PR00010; EGFBLOOD.
DR PRINTS: PR00011; EGFLAMININ.
DR PRINTS: PR00012; ENTPEI.
DR PRINTS: PR00286; CHARYBDOTOXIN.
DR PRINTS: PR00287; THIONIN.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; UNKNOWN_7.
DR PROSITE: PS01186; EGF_2; 6.
DR PROSITE: PS01187; EGF_CA; 4.
KW Glycoprotein; EGF-like domain.
FT NON_TER 1
FT NON_TER 263
FT SEQUENCE 263 AA; 28140 MW; 971C7E61C2B7F7B8 CRC64;
Query Match 8.9%; Score 202.5; DB 4; Length 263;
Best Local Similarity 44.94; Pred. No. 2.3e-10;
Matches 35; Conservative 5; Mismatches 31; Indels 7; Gaps 1;
QY 3 CTNPCHNGCCEISEAVRQDTFTGVCKPRFGHICQHNINEGEVERCKNGGICDL 62
DB 25 CTSSCFNGGV-----DGINSFCLCPVFTGSCFLHEINCSHPCLNDGTCTVCG 77
QY 63 VANYSCCEGCEMGRNQ 80
DB 78 LGTYRSCPLGTGTRNCO 95
RESULT 9
ID Q99740 PRELIMINARY; PRT; 263 AA.
AC Q99740:
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE SOLUBLE PROTEIN JAGGED (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL VEIN;
RX MEDLINE=95211842; PubMed=7697721;
RA Lindell C.E., Shaver C.J., Boulter J., Weinmaster G.;
RT "Jagged: a mammalian ligand that activates Notch1."
RL Cell 80:909-917(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=UMBILICAL VEIN;
RX MEDLINE=97115768; PubMed=8955070;
RA Zimin A.B., Pepper M.S., McMahon G., Nguyen F., Montesano R.,
RA Maciag A.;
RT "An antisense oligonucleotide to the notch ligand jagged enhances
RT fibroblast growth factor-induced angiogenesis in vitro."
RL J. Biol. Chem. 271:32499-32502(1996).
DR EMBL: U77914; AAC50909.1; -
DR HSSP: P00743; ICCF.
DR INTERPRO: IPR000083; -
DR INTERPRO: IPR000152; -
DR INTERPRO: IPR000561; -
DR INTERPRO: IPR000742; -
DR INTERPRO: IPR001881; -

```

DR INTERPRO: IPR002049; -
 DR PFAM: PF00008; EGF_3;
 DR PRINTS: PRO0011; EGF_LAMININ.
 DR PRINTS: PRO0012; ENTPEP.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS01186; EGF_2; 2;
 DR PROSITE: PS01187; EGF_CA; 2;
 KM Glycoprotein; EGF-like domain.
 FT NON_TER 1
 SQ SEQUENCE 263 AA; 28405 MW; E77235CB360B620 CRC64;

Query Match 8.8%; Score 199.5; DB 4; Length 263;
 Best Local Similarity 32.5%; Pred. No. 4.3e-10;
 Matches 40; Conservative 8; Mismatches 44; Indels 31; Gaps 2;

OY 3 CTRPCHNGTCEIS-----EAYRGDTFIG---YVCK 31
 |||||
 DB 6 CLPNPCHNGTGVNGESFTCVCKEGWESPTCAQNTNDCSPHPCVNGSTGVGDGMWYRCE 65
 OY 32 CPGSGNGJHCHNINCEVEPCCKNGICTDLVANSCECPGEGFMGRNGCYKCSGLGIEG 91
 |||||
 DB 66 CAPEFAPPCRIINCCSPAPAFATCVPEINICVCPCHSGAKQCEVSGRCITMG 125
 OY 92 GIT 94
 :
 DB 126 SVI 128

RESULT 10
 006007 PRELIMINARY; PRT; 387 AA.

ID 006007;
 AC 006007;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE NOTCH PROTEIN HOMOLOG 1 (NOTCH A PROTEIN) (FRAGMENT).
 GN NOTCH OR NOTCH A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBA X C57BL; TISSUE=EMBRYO;
 RX MEDLINE=93178563; PubMed=8440332;
 RA Lardelli M., Lendahl U.;
 RT "Notch A and notch B--two mouse Notch homologues coexpressed in a wide
 variety of tissues";
 RT Exp. Cell Res. 204:364-372(1993).
 DR EMBL: X68278; CAA8339.1; -;
 DR HSP; P00743; ICGF.
 DR MD; MGI:97363; Notch1.
 DR INTERPRO: IPR000083; -;
 DR INTERPRO: IPR000152; -;
 DR INTERPRO: IPR000561; -;
 DR INTERPRO: IPR000742; -;
 DR INTERPRO: IPR000800; -;
 DR INTERPRO: IPR001438; -;
 DR INTERPRO: IPR001881; -;
 DR INTERPRO: IPR002049; -;
 DR PFAM: PF00008; notch; 6;
 DR PFAM: PF00066; notch; 3;
 DR PRINTS: PRO0010; EGF_LAMININ.
 DR PRINTS: PRO0011; EGF_LAMININ.
 DR PRINTS: PRO0012; ENTPEP.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_7.
 DR PROSITE: PS01186; EGF_2; 3;
 DR PROSITE: PS01187; EGF_CA; 2;
 KW Differentiation; Neurogenesis; Repeat.
 FT NON_TER 1
 SQ SEQUENCE 387 AA; 41497 MW; 6E82AC932602D449 CRC64;

FT NON_TER 387
 SQ SEQUENCE 387 AA; 41497 MW; 6E82AC932602D449 CRC64;

Query Match 8.7%; Score 197.5; DB 11; Length 387;
 Best Local Similarity 37.9%; Pred. No. 1.1e-09;
 Matches 33; Conservative 15; Mismatches 22; Indels 17; Gaps 4;

OY 3 CTRPCHNGTCEISAYRGDTFIGYCKCPRGFGHCHNINCE--VEP-----C 53
 |||||
 DB 27 CLSOPCHNGTCTDLINISX-----CSPRGTYGKEINVDCHPPLDPAASRPRK 78
 OY 54 KNGICTDLVANSCECPGEGFMGRNQ 80
 |||||
 DB 79 FNGTGVGVGVGYCTCPGPEVERGE 105

RESULT 11
 046370 PRELIMINARY; PRT; 308 AA.

ID 046370;
 AC 046370;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE PRELIMINARY FACTOR-1.
 GN DLK.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FAF;
 RA Minoshima Y., Taniguchi Y., Sasaki Y.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Fahrenkrug S.C., Freking B.A., Smith T.P.S.;
 RT "Genomic Organization and Chromosomal Position of the Bovine DLK
 Gene";
 RT EMBL; AB009278; BAA2373.1; -;
 DR EMBL; AB009278; BAA2373.1; -;
 DR EMBL; AF181466; AF00926.1; -;
 DR EMBL; AF181463; AF00926.1; JOINED.
 DR EMBL; AF181464; AF00926.1; JOINED.
 DR EMBL; AF181465; AF00926.1; JOINED.
 DR EMBL; AF181462; AAD56944.1; -;
 DR HSP; P35555; IEMN.
 DR INTERPRO: IPR000152; -;
 DR INTERPRO: IPR000561; -;
 DR INTERPRO: IPR000742; -;
 DR PFAM; PF00008; EGF_5.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_4.
 DR PROSITE: PS01186; EGF_2; 5;
 KW Glycoprotein.
 SQ SEQUENCE 308 AA; 33005 MW; FF42B9103D9D852 CRC64;

Query Match 8.4%; Score 189.5; DB 6; Length 308;
 Best Local Similarity 38.1%; Pred. No. 4.2e-09;
 Matches 37; Conservative 11; Mismatches 40; Indels 9; Gaps 3;

OY 1 GCTPN--PCNNGTCEISAYRGDTFIGYCKCPRGFGHCHNINCEVEPCCKNGI 58
 |||||
 DB 129 GTCVYNGSPCHGSGCVDDGGRAPHA----VLCPRGSSNGFCIVTNSIPRNCNGI 184
 OY 59 CTDLVANSCECPGEGFMGRNQ---YKCSGLGIEG 92
 |||||
 DB 185 CTDLGDFRCRCPAGFMDKCSRPVNTCTSEPCING 221

FT NON_TER 192 192
SQ SEQUENCE 192 AA: 20285 MM: 34BAD3011C30AB3 CRC64;
Query Match 8.1%; Score 182.5; DB 4; Length 192;
Best Local Similarity 31.2%; Pred. No. 9.5e-09;
Matches 35; Conservative 11; Mismatches 35; Indels 31; Gaps 3;
QY 2 PCTPNCHNGGTCEI-----SEAYRDTF-----IGYVCK 31
DB 50 PCASSPCQNNGTCTVDYVHTFCNCSPEFTGPTCAOLIDPCALSPCAHGTCSRVSSTYKCL 109
QY 32 CPFGFNGHCHNINECEVEPCCKNGGICIDLVANYSCEGPEFGNCO-YK 82
DB 110 CDPGYNGICCEFTNECISAPCLMAATCRDLVNGYECYCLAFYKTHCELYK 161
RESULT 15
Q09PK6 PRELIMINARY; PRT; 315 AA.
AC Q09PK6;
DT 01-MAY-2000 (TREMBlrel.13, Created)
DT 01-MAY-2000 (TREMBlrel.13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel.15, Last annotation update)
DE PGCN_HUMAN, PARTIAL CDS (FRAGMENT).
GN NEUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lameirin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., StiWagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Dangnan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Altix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommshli A., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kopyayshi A., Olsen A.S., Carrano A.V.;
PT "Sequence analysis of an ~1 Mb region containing the MEF2B gene in
PT 19p12."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005254; AAC25581.1;
DR HSE: P00740; IIXA.
DR INTERPRO: IPR00132;
DR INTERPRO: IPR00436;
DR INTERPRO: IPR00361;
DR INTERPRO: IPR00742;
DR INTERPRO: IPR01010;
DR INTERPRO: IPR01304;
DR INTERPRO: IPR01438;
DR INTERPRO: IPR01881;
DR INTERPRO: IPR01947;
DR PFAM: PF00008; EGF_2.
DR PFAM: PF00059; lectin_C; 1.
DR PFAM: PF00084; sushi_1.
DR PRINTS: PRO0010; EGFBLD.
DR PRINTS: PRO0286; CHARYBDTOXIN.
DR PRINTS: PRO0287; THIONIN.
DR PRINTS: PRO0356; ANTIFREEZEII.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 315 AA: 36306 MM: 1813208897ED4AB CRC64;

Query Match 8.1%; Score 182.5; DB 4; Length 315;
Best Local Similarity 26.9%; Pred. No. 1.8e-08;

Matches 54; Conservative 24; Mismatches 62; Indels 61; Gaps 9;
QY 2 PCTPNCHNGGTCEI-----SEAYRDTF-----IGYVCKCPRGNGIHCOHINIECEVEPCCKNGGICTD 61
DB 5 PCNNPCLHGTGTCNNANGTM-----YGCSCDQGFAGENCEDIDDLCSPEENGSTCID 57
QY 62 LVANYSCEGPEFGNRCQYKCSGPIGIGIISNQITASSSTHRALEGLQW---YPY 117
DB 58 EVNGFVCLCLPSYGSPCE-----KDEGCDRGWHRFGGHCYRY 96
QY 118 YARLNKKGILINMTAENDRWKRWIGINLORRMVTVITQAGARISPEYIKFYKIAS 177
DB 97 FAHR-----RAMEDAEKDCRR-----SGHLS-----VNSPEHSFTN-SFG 133
QY 178 NDGKTWAMKYKVGTEMDYFR 198
DB 134 HE-NTWI-----GLNDRIYER 148

Search completed: May 23, 2001, 08:37:26
Job time: 127 sec

PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Example; Column 63-64; 73pp; English.
 CC

CC The present sequence represents human milk fat globule protein MFG-E8,
 CC which has homology to the developmentally-regulated endothelial cell
 CC locus 1 (Del-1). The Del-1 protein has epidermal growth factor (EGF) like
 CC domains and discoidin-1/factor VIII-like domains. The Del-1 proteins
 CC have an inhibitory effect on angiogenesis (blood vessel growth), this
 CC activity may be useful clinically to prevent neovascularization of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC Integrin, and is an apoptosis inducer.
 CC
 XX Sequence 321 AA;

Query Match 75.8%; Score 1717; DB 20; Length 321;
 Best Local Similarity 98.4%; Pred. No. 1,2e-110;
 Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 83 CSGPLGEGGIIISNOQITASTSTRALFGLOKVPYARLNKGLINAWTAENDRKRMI 142
 DB 1 csqpllgiegllsnqqltassthralfgkvypparlnkglinawtaendrwrvl 60
 QY 143 QINQRMRYTGITOGAKRIGSPYIKFYKINYSNDGKTAMKVKVGTEDVPRGND 202
 DB 61 qinqrmyrtgvtogakrigspykfykiklaysndgktamkvkvgtedmvfrgnd 120
 QY 203 NNPFYASFPPIKAQVRLPYOCRRHCTLRMELIGCTSCGSPRIGMSGHIQYOT 262
 DB 121 nnpfyastfpkqgyrlypyccrrhctlmellgctscgsepriymsgnldqyqlt 180
 QY 263 ASSTFRLNMDFTWEPKARLDKOGKVNMTSGHNDOSOMLOVDLYPRKYVCTTQCA 322
 DB 181 asstfllnmdfwepkardkqgvnmwtsghndqsgqlqyxllyprkvglltqga 240
 QY 323 KDRCHQVFGSYKILAYSDNGEHMTVODEKQKDKYFGCNEEDNTHRRKNVIDPPIARI 382
 DB 241 kdxhvgfvsykyklaysndgenwcvxqdekqkdkvkgndndhrrknvidppiyarh 300
 QY 383 RILPMSYGRITLASELLGCT 403
 DB 301 rilpmsygritlaseillgct 321

RESULT 2
 R77252 R77252 standard; Protein; 387 AA.
 AC R77252;
 XX 21-NOV-1995 (first entry)
 DT
 XX HMEG 46 kDa antigen.
 DE
 XX HMEG; human milk fat globule; antigen; immunogen; vaccine; virucide;
 KW epithelium; tumor; breast cancer; monoclonal antibody; Mab.
 XX Homo sapiens.
 OS
 XX MO9515171-A.
 PN
 XX 08-JUN-1995.
 PD
 XX 05-DEC-1994. 94WO-US13967.
 PF
 XX 03-DEC-1993. 93US-0162402.
 PR

XX (CANC-) CANCER RES FUND CONTRA COSTA.
 PA Certani RL, Larocca DJ, Peterson JA;
 XX WPI, 1995-215151/28.
 XX DR N-PSDB; Q91198.
 XX
 XX 46 kD apparent molecular weight human milk fat globule antigen
 PT used in assays to determine the presence of a cancerous tumour of
 PT epithelial origin, and in a vaccine against neoplastic tumours
 CC
 XX Claim 6; Page 46-47; 68pp; English.
 CC
 CC A complete cDNA sequence for the 46 kDa HMEG antigen, a major
 CC component of the apical surface of the normal breast epithelial
 CC cell, was obtained by PCR and RACE methods. cDNA clones can be
 CC used to prepare MAbs for use in immunotherapy, immunohistochemistry,
 CC prognosis, diagnosis, imaging and therapy. Recombinant antigen can
 CC be expressed in prokaryotic or (glycosylated) in eukaryotic cells.
 CC
 XX Sequence 387 AA;

Query Match 45.1%; Score 1020.5; DB 16; Length 387;
 Best Local Similarity 48.8%; Pred. No. 1e-62;
 Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;
 QY 3 CTNPCHNGGTC-EISEAYRGDFIGYCKCPRGFNGHCHGNHNECEVPCNGSGICTD 61
 DB 27 csknpnhnggleelsgevgvdfvpytctclkyganh----- 65
 QY 62 LVANVCECEPGFMRNOCYKSGPLGEGGIIISNOQITASTSTRALFGLOKVPYARL 121
 DB 66 -----celkveplgmenqnlansqlaasvvtflgqmwpealari 108
 QY 122 NKGLINAWTAENDRKRMIQINLQKRMRYGVITOGAKRIGSPYIKFYKINYSNDG 181
 DB 109 nrgmwnawcpesnd nptqynllrmvvgvvgqastlaahyikaikvayslng 167
 QY 182 TW-AMKVKGINMEDVFRGNIDNNTPYASFPPIKAQVRLPYOCRRHCTLRMELIG 240
 DB 168 etafllndvknkhefv- gmnknavhnlfelvpeagvlylpschactclrfellgc 225
 QY 241 ELSGCEPLGMSGHIQYOTIYASSTFRLNMDFTWEPKARLDKOGKVNMTSGHNDQ 300
 DB 226 elngcanpqlglnmslpdkqltasssyktwgllfswmpsyarldkqgnfnawagaygn 285
 QY 301 SQMLQVDLYPRKYVCTTQCAKDPGHVQVFGSYKILAYSDNGEHMTVODEKQKDKYFO 360
 DB 286 dqwlgvdlgsakevltgltgarnfgsvqfvasykvaysndsanwteygdprtgasklp 345
 QY 361 GNFDNDTHRRKNVIDPPIYARIRILPMSYGRITLASELLGCT 402
 DB 346 gnvndshkknlfecplaryvrlpwahnzlarlleiigc 387

RESULT 3
 Y94453 Y94453 standard; Protein; 387 AA.
 AC Y94453;
 XX 11-SEP-2000 (first entry)
 DT
 XX Human lactadherin protein.
 DE
 XX Human, lactadherin; MCF-E8; anti-tumour; immune response;
 KW exosome; dendritic cell.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 PR


```

XX      Sequence      218 AA;
SQ
Query Match:          28.4%; Score 643; DB 16; Length 218;
Best Local Similarity 56.8%; Pred. No. 4,5e-37;
Matches 117; Conservative 30; Mismatches 59; Indels 0; Gaps 0;

OY 197 FRGNDINMPFANSFTPTPIKAOYRLYPOVCRHCTRMELLGCELCSCSEPLGMSGH 256
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 13 fvgwnknavhvnlfecpveagvyrlyptschactrlfcllgycelngcamlgiknmsl 72

OY 257 ODYOTASSIFRTLMNMFMEPRKARLDKOGKVNMTSGHNDQOMLVNDLVPTRKYG 316
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 73 pdqqltaassykktwglhlfswpsyarldkgnfnawvagsygnqqlvqlgsksevg 132

OY 317 IIFOGAKDFGVHOFVSGYKLAYSNDGEMHTVYODEKOROKYFOGNFNDTHRKNTVDP 376
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 133 lltqsrnlfgsqvfvasykwaynsdnamteygdqprtgsaklfpgnwhshkknlfecp 192

OY 377 IYARHRIILPMSWYGRITLASRLIGC 402
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 193 llaryvrllypawhmltalrlcllgc 218

RESULT 9
W10367
ID W10367 standard; Protein: 85 AA.
XX
AC W10367;
XX
DT 03-MAY-1997 (first entry)
XX
DE Del-1 discoidin I-like domain.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1;
KW signal transduction; cancer; tumour marker; angiogenesis;
KW diabetic retinopathy; rheumatoid arthritis; endometriosi;
KW cardiac ischaemia; stroke; vascular disease; wound healing;
KW vulnerability; bone formation; diagnosis; therapy; discoidin.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 16 /note= "unspecified amino acid residue"
FT Misc-difference 17 /note= "unspecified amino acid residue"
FT Misc-difference 84 /note= "unspecified amino acid residue"
FT Misc-difference 85 /note= "unspecified amino acid residue"
FM Misc-difference 85 /note= "unspecified amino acid residue"
XX
FN W09640769-A1.
XX
PD 19-DEC-1996.
XX
PE 05-JUN-1996; 96MO-US09456.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B. Quettermous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI; 1997-052233/05.
XX
New developmentally regulated endothelial cell locus-1 (del-1) gene
XX - used to develop prods. for the diagnosis and treatment of cancer
XX and conditions involving abnormal angiogenesis
XX
PS Example: Page 71-72; 137pp; English.

```

```

XX      A discoidin-like domain (W10367) was identified in mouse
CC developmentally-regulated endothelial cell locus-1 (Del-1)
CC sequences (see also W10364) derived from a trapped exon and
CC mouse embryo cDNAs. It showed homology to discoidin-like domains
CC from human milk fat globule protein, human factor V, mouse factor
CC VIII, X-A5b1 and X-A5b2 (domains of Xenopus neuronal antigen A5)
CC and discoidin I.
XX
SQ      Sequence      85 AA;
SQ
Query Match:          17.1%; Score 387; DB 18; Length 85;
Best Local Similarity 91.5%; Pred. No. 6e-20; 4; Indels 2; Gaps 1;
Matches 75; Conservative 1; Mismatches 4;

OY 307 DLIVPTKVTGILITOG--AKDFGVHOFVSGYKLAYSNDGEMHTVYODEKOROKYFOGNF 364
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 dllypkrvtglltqgxakdldgvlfvgsyklaysndgenmwnvndekqtkdvfgnid 60

OY 365 NDRHKNVDPPIYARHRIILP 386
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ndthknvldpplyarfilp 82

RESULT 10
W94689
ID W94689 standard; Peptide: 85 AA.
XX
AC W94689;
XX
DT 05-MAY-1999 (first entry)
XX
DE Discoidin I/factor VIII like domain #1.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosi;
KW angiogenesis.
XX
OS Mus sp.
XX
FH Synthetic.
XX
PD US5877281-A.
XX
PE 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1995; 96US-0659235.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B. Quettermous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI; 1999-189720/16.
XX
Del-1 polypeptide sequences - useful for treatment of cancer,
XX diabetic retinopathy, rheumatoid arthritis and endometriosi
XX
Example: Column 37-38; 73pp; English.
XX
CC The present sequence represents a discoidin I/factor VIII-like domain.
CC The present invention describes developmentally-regulated endothelial
CC cell locus 1 (del-1). Del-1 has epidermal growth factor like domains and
CC inhibitory effect on angiogenesis (blood vessel growth), this activity
CC may be useful clinically to prevent neovascularisation of tissues such
CC as tumour nodules and prevention of metastases. The anti-angiogenic
CC activity of Del-1 may be used to treat abnormal conditions that result
CC from angiogenesis, including cancer, diabetic retinopathy, rheumatoid

```


CC arthritis and endometriosi. Since Del-1 promotes angiogenesis it can be
 CC used to treat cardiac ischemia, thrombotic stroke, wound healing and
 CC peripheral vascular disease. Del-1 is also useful for promoting bone
 CC formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis
 CC inducer.
 CC
 SQ Sequence 85 AA;

Query Match 17.1%; Score 387; DB 20; Length 85;
 Best Local Similarity 91.5%; Pred. No. 66-20;
 Matches 75; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

OY 307 DLLVPTKVTGTTTGG--AKDFGHVQFVGSYKLAISNDGEHMTVYQDEKORRKFVPGNPD 364
 DB 1 dllyptkvtglltggxxakdfghvfygsklaynsdgehmtvhdqekpfkdvfgnld 60
 OY 365 NDTRRKNVDPPIYARHRIIP 386
 DB 61 ndthknvdppliyarfirllp 82

RESULT 11
 ID W10370 standard; Protein; 102 AA.
 AC W10370;
 DT 03-MAY-1997 (first entry)
 DE Human Del-1 splice variant partial sequence.

KM Del-1; developmentally-regulated endothelial cell locus-1;
 KM signal transduction; cancer; tumor marker; angiogenesis;
 KM diabetic retinopathy; rheumatoid arthritis; endometriosi;
 KM cardiac ischaemia; stroke; vascular disease; wound healing;
 KM vulnery; bone formation; diagnosis; therapy.

OS Homo sapiens.
 PH Key Location/Qualifiers
 FT 3..36 /label= EGF-1
 FT /note= "epidermal growth factor-like domain 1"
 FT 45..83 /label= EGF-2
 FT /note= "epidermal growth factor-like domain 2"
 FT 90..102 /label= EGF-3
 FT /note= "N-terminal portion of epidermal growth
 factor-like domain 3"

PN W09640769-A1.
 PD 19-DEC-1996.
 PE 05-JUN-1996; 96MO-US09456.
 PR 07-JUN-1995; 95US-0480229.
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 PI Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
 DR WPI: 1997-052233/05.
 XX New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 CC and conditions involving abnormal angiogenesis
 XX Example; Page 91: 137bp; English.

CC A polypeptide sequence (W10370) shows a portion of a variant form
 CC of human developmentally-regulated endothelial cell locus-1 (del-1)
 CC polypeptide. In comparison with the major form (W10365) of human
 CC Del-1, 10 amino acid residues are missing between EGF-like domains
 CC 1 and 2 of the Del-1 variant. This is a result of alternative
 CC splicing (see also T47340).
 CC
 SQ Sequence 102 AA;

Query Match 16.3%; Score 368; DB 18; Length 102;
 Best Local Similarity 100.0%; Pred. No. 1.5e-18;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GPCPTNPCHNGCTEISEAVRGDTFYGVCNCPRGNGIHCOHNIKEVEPCKNIGICT 60
 DB 43 gpcptnpchngctcseavrygdtfygvcnkpgrngfhqgmnecevepknngict 102

RESULT 12
 ID W69527 standard; peptide; 160 AA.
 AC W69527;
 DT 09-NOV-1998 (first entry)
 DE C2 peptide SEQ ID NO:9 from W09829453 Example 3.

KM Drug; cell membrane-directed drug; phospholipid; lipid bilayer;
 KM cell cortex; blood coagulation; inflammation; immunological disorder.

OS Synthetic.
 OS Homo sapiens.
 PN W09829453-A1.
 PD 09-JUL-1998.

PE 05-JAN-1998; 98MO-JP00002.
 PR 27-DEC-1996; 96JP-0359053.
 PA (MOCH) MOCHIDA PHARM CO LTD.
 PI Hasegawa T, Kuriyama S;

DR WPI: 1998-388051/33.
 DR N-PSDB; VA0051.

PT Drugs containing peptide(s) with specific affinity to
 PT phospholipid(s) - such as phosphatidyl serine, for treatment of
 PT blood coagulation, inflammatory and immunological disorders
 XX
 XX Example 3; Page 72-73; 117bp; Japanese.

CC The present invention describes drug compositions which contain as
 CC an active component a peptide which has specific affinity to particular
 CC phospholipids (such as phosphatidyl ethanolamine or phosphatidyl serine),
 CC especially to phospholipids which constitute a lipid bilayer of cellular
 CC cortex and of which the concentration in the bilayer increases in cells
 CC which are abnormal (e.g. through injury, denaturation or activation). In
 CC particular, the peptide contains a sequence having phospholipid affinity
 CC and a structure of formula (I): (A1)a-(A2)b-(A3)c, where (A1) is one of
 CC two specific sequences (see W69516 and W69519), (A2) and (A3) are
 CC TYRIRHPOSMVHQIALR, LRYRIRHPOSMVHQIALR (see W69517) or MEVIGCEAQNLY (see
 CC W69518); a = 0-5; b = 1-5, and c = 0-5. Preferred are the formulae
 CC A1-A2-A3, A2-A3, A2-A2-A3 or A2-A2 (especially A2-A2-A3,
 CC A2-A2-A3 or A2-A2). The sequence is linked to a peptide such as a
 CC blood factor, especially thrombo-modulin, urina-statin or membrane
 CC cofactor protein. The drugs are used for the treatment and prevention of
 CC diseases involving blood coagulation, inflammatory and immunological
 CC disorders. The present sequence represents a peptide from the present

CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC Integrin, and is an apoptosis inducer.

XX Sequence 43 AA;

Query Match 11.7%; Score 264; DB 20; Length 43;

Best Local Similarity 100.0%; Pred. No. 8e-12;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CPNPGHNGGCGEISEAYRGDTFISGVCKCPRGNGHCHQNI 45
 |||

Db 1 clpncphngggtceiseayrgdtfisyvckcprgfnghcqhni 43

RESULT 15

W94690
 ID W94690 standard; peptide: 85 AA.

XX W94690;

DT 05-MAY-1999 (first entry)

DE Discoidin I/Factor VIII like domain #2.

XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;

KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;

KW diabetic retinopathy; rheumatoid arthritis; endometriosis;

XX anglogenesis.

OS Homo sapiens.

XX Synthetic.

XX US5877281-A.

XX 02-MAR-1999.

XX 05-JUN-1996; 96US-0659235.

XX 05-JUN-1996; 96US-0659235.

XX 07-JUN-1995; 95US-0480229.

XX (PROG-) PROGENITOR INC.

XX (UYVA-) UNIV VANDERBILT.

XX Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;

XX WPI: 1999-189720/16.

XX Del-1 polypeptide sequences - useful for treatment of cancer,

XX diabetic retinopathy, rheumatoid arthritis and endometriosis

XX Example: Column 39-40; 73pp; English.

XX The present sequence represents a discoidin I/factor VIII-like domain.

XX The present invention describes developmentally-regulated endothelial

XX cell locus 1 (Del-1). Del-1 has epidermal growth factor like domains and

XX discoidin I/factor VIII-like domains. The Del-1 proteins have an

XX inhibitory effect on angiogenesis (blood vessel growth), this activity

XX may be useful clinically to prevent neovascularisation of tissues such

XX as tumour nodules and prevention of metastases. The anti-angiogenic

XX activity of Del-1 may be used to treat abnormal conditions that result

XX from angiogenesis, including cancer, diabetic retinopathy, rheumatoid

XX arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be

XX used to treat cardiac ischaemia, thrombotic stroke, wound healing and

XX peripheral vascular disease. Del-1 is also useful for promoting bone

XX formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis

XX inducer.

XX Sequence 85 AA;

XX Query Match 11.1%; Score 252; DB 20; Length 85;

Best Local Similarity 57.3%; Pred. No. 1.1e-10;

Matches 47; Conservative 14; Mismatches 19; Indels 2; Gaps 1;

QY 307 DLIVPTKVTGIIITG--AKDEGHVQYVGSIKLAYSNDGEHNTVYQDEKQKRDVFGQND 364
 |||

Db 1 dlgskevgtlltgxxarnlgsvglvasykvayndsantleyqdprrtgsksvrtgnd 60

QY 365 NDTNRKNVIDPEIVARHTRILP 386
 |||

Db 61 nshkknlfexpfmaryrvrlp 82

Search completed: May 23, 2001, 08:35:17
 Job time: 108 sec

Wed May 23 08:50:18 2001

us-09-237-981-14_copy_108_513.clo.ram

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:34:44 ; Search time 45.61 Seconds
(without alignments)
1431.856 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Perfect score: 2264

Sequence: 1 GPCPNPCNNGSTCEISEAV.....MSWGRITLASELIGCTEEK 406

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 903930

Minimum DB seq length: 0

Maximum DB seq length: 406

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending Patents, AA_Main:*

1: /cgn2_6/prodata/2/paa/US087.COMB.pep.*
2: /cgn2_6/prodata/2/paa/US088.COMB.pep.*
3: /cgn2_6/prodata/2/paa/US089.COMB.pep.*
4: /cgn2_6/prodata/2/paa/US090.COMB.pep.*
5: /cgn2_6/prodata/2/paa/US091.COMB.pep.*
6: /cgn2_6/prodata/2/paa/US092.COMB.pep.*
7: /cgn2_6/prodata/2/paa/US093.COMB.pep.*
8: /cgn2_6/prodata/2/paa/US094.COMB.pep.*
9: /cgn2_6/prodata/2/paa/US095.COMB.pep.*
10: /cgn2_6/prodata/2/paa/US096.COMB.pep.*
11: /cgn2_6/prodata/2/paa/US097.COMB.pep.*
12: /cgn2_6/prodata/2/paa/US098.COMB.pep.*
13: /cgn2_6/prodata/2/paa/US099.COMB.pep.*
14: /cgn2_6/prodata/2/paa/US100.COMB.pep.*
15: /cgn2_6/prodata/2/paa/US101.COMB.pep.*
16: /cgn2_6/prodata/2/paa/US102.COMB.pep.*
17: /cgn2_6/prodata/2/paa/US103.COMB.pep.*
18: /cgn2_6/prodata/2/paa/US104.COMB.pep.*
19: /cgn2_6/prodata/2/paa/US105.COMB.pep.*
20: /cgn2_6/prodata/2/paa/US106.COMB.pep.*
21: /cgn2_6/prodata/2/paa/US107.COMB.pep.*
22: /cgn2_6/prodata/2/paa/US108.COMB.pep.*
23: /cgn2_6/prodata/2/paa/US109.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1717	75.8	321	US-08-480-229B-21	Sequence 21, Appl
2	1717	75.8	321	US-08-659-235B-21	Sequence 21, Appl
3	1717	75.8	321	US-09-237-981-21	Sequence 21, Appl
4	1405.5	62.1	362	US-60-207-315-523	Sequence 523, App
5	1405.5	62.1	311	US-60-230-435-1519	Sequence 1519, Ap
6	1020.5	45.1	387	US-09-582-340-2	Sequence 2, Appl
7	948	41.9	320	US-08-480-229B-20	Sequence 20, Appl
8	948	41.9	320	US-08-659-235B-20	Sequence 20, Appl
9	948	41.9	320	US-09-237-981-20	Sequence 20, Appl
10	816.5	36.1	203	US-60-233-644-77	Sequence 77, Appl

11	813.5	35.9	185	US-60-230-435-1806	Sequence 1806, Ap
12	784.5	34.7	221	US-09-237-981-29	Sequence 29, Appl
13	643	28.4	217	US-08-482-596A-3	Sequence 3, Appl
14	643	28.4	218	US-08-482-596A-2	Sequence 2, Appl
15	494	21.8	218	US-08-482-596A-4	Sequence 4, Appl
16	475.5	21.0	218	US-08-482-596A-5	Sequence 5, Appl
17	387	17.1	85	US-08-480-229B-1	Sequence 1, Appl
18	387	17.1	85	US-08-659-235B-1	Sequence 1, Appl
19	387	17.1	85	US-09-237-981-1	Sequence 1, Appl
20	373	16.5	159	PCT-US00-22226A-11	Sequence 11, Appl
21	373	16.5	159	PCT-US00-22226A-11	Sequence 11, Appl
22	368	16.3	109	US-08-463-294-25	Sequence 25, Appl
23	364.5	16.1	160	PCT-US00-22226-8	Sequence 8, Appl
24	364.5	16.1	160	PCT-US00-22226A-8	Sequence 8, Appl
25	351	15.5	161	US-60-209-043-215	Sequence 165, App
26	351	15.5	161	US-60-209-043-215	Sequence 215, App
27	344.5	15.2	160	US-09-331-793-9	Sequence 9, Appl
28	343.5	15.2	160	PCT-US00-22226A-9	Sequence 9, Appl
29	343.5	15.2	160	PCT-US00-22226A-9	Sequence 9, Appl
30	335.5	14.8	160	PCT-US00-22226-10	Sequence 10, Appl
31	335.5	14.8	160	PCT-US00-22226A-10	Sequence 10, Appl
32	323	14.3	150	PCT-US00-22226-12	Sequence 12, Appl
33	323	14.3	150	PCT-US00-22226A-12	Sequence 12, Appl
34	319.5	14.1	62	US-60-192-739-1075	Sequence 165, App
35	317	14.0	126	US-60-233-940-165	Sequence 165, App
36	304	13.4	110	US-08-463-294-24	Sequence 24, Appl
37	301.5	13.3	96	US-60-169-840-5111	Sequence 5111, Ap
38	291.5	13.3	96	US-60-169-867-4347	Sequence 4347, Ap
39	295.5	13.1	56	US-60-171-489-1052	Sequence 1052, Ap
40	287.5	12.7	160	PCT-US00-22226-7	Sequence 7, Appl
41	287.5	12.7	160	PCT-US00-22226A-7	Sequence 7, Appl
42	287.5	12.7	390	US-60-207-315-488	Sequence 488, App
43	280.5	12.3	112	US-60-230-435-1526	Sequence 1526, Ap
44	279.5	12.3	130	US-09-331-793-1	Sequence 1, Appl
45	278	12.3	52	US-60-236-359-21193	Sequence 21193, A

ALIGNMENTS

RESULT 1
US-08-480-229B-21
Sequence 21, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480, 229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Polsant, Brian M.
REGISTRATION NUMBER: 28, 462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
 TELE: 66141 Pennie
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 321 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-480-229B-21

Query Match 75.8%; Score 1717; DB 8; Length 321;
 Best Local Similarity 98.4%; Pred. No. 5.6e-151;
 Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 83 CSGPLGIEGGIISNOQITASSTRHALFGLQKWPYVARLNKKGLINMTAENDRRKRWI 142
 DB 1 CSGPLGIEGGIISNOQITASSTRHALFGLQKWPYVARLNKKGLINMTAENDRRRWI 60

QY 143 QINLQRRKRVGTITOGAKRIGSPEYIKFYKTIAYSDGKTWAMYKVGTEDEWFRGNID 202
 DB 61 QINLQRRKRVGTITOGAKRIGSPEYIKFYKTIAYSDGKTWAMYKVGTEDEWFRGNID 120

QY 203 NNTPYANSFTPIKQOYRILYPOYCRHRCTLRMELLGCELSGCSPEPLGMSGHIQOYQIT 262
 DB 121 NNTPYANSFTPIKQOYRILYPOYCRHRCTLRMELLGCELSGCSPEPLGMSGHIQOYQIT 180

QY 263 ASSIFFTLMDMFTWEPKARLDKQGVNMTSGHNDOSOMLOYDLLVPTKVTGIIITOGA 322
 DB 181 ASSIFFTLMDMFTWEPKARLDKQGVNMTSGHNDOSOMLOYDLLVPTKVTGIIITOGA 240

QY 323 KDGCHQVFGSKTIAYSDGGEHMTVYODEKORKDKVPGCNPDNTHKKNYIDPPIYARIH 382
 DB 241 KDGCHQVFGSKTIAYSDGGEHMTVYODEKORKDKVPGCNPDNTHKKNYIDPPIYARIH 300

QY 383 RILPMSWYGRITLASSELLGCT 403
 DB 301 RILPMSWYGRITLASSELLGCT 321

RESULT 2
 US-08-659-235B-21
 ; Sequence 21, Application US/08659235B
 ; GENERAL INFORMATION:
 ; APPLICANT: Quettermous, Thomas
 ; APPLICANT: Hogan, Bridgid
 ; APPLICANT: Snodgrass, H. Ralph
 ; APPLICANT: Zupancic, Thomas J.
 ; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 ; TITLE OF INVENTION: CELL LOCUS-1
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/659,235B
 ; FILING DATE: 05-JUN-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Poissant, Brian M.
 ; REGISTRATION NUMBER: 28,462
 ; REFERENCE/DOCKET NUMBER: 8907-034
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELE: 66141 Pennie
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 321 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-659-235B-21

Query Match 75.8%; Score 1717; DB 10; Length 321;
 Best Local Similarity 98.4%; Pred. No. 5.6e-151;
 Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 83 CSGPLGIEGGIISNOQITASSTRHALFGLQKWPYVARLNKKGLINMTAENDRRKRWI 142
 DB 1 CSGPLGIEGGIISNOQITASSTRHALFGLQKWPYVARLNKKGLINMTAENDRRRWI 60

QY 143 QINLQRRKRVGTITOGAKRIGSPEYIKFYKTIAYSDGKTWAMYKVGTEDEWFRGNID 202
 DB 61 QINLQRRKRVGTITOGAKRIGSPEYIKFYKTIAYSDGKTWAMYKVGTEDEWFRGNID 120

QY 203 NNTPYANSFTPIKQOYRILYPOYCRHRCTLRMELLGCELSGCSPEPLGMSGHIQOYQIT 262
 DB 121 NNTPYANSFTPIKQOYRILYPOYCRHRCTLRMELLGCELSGCSPEPLGMSGHIQOYQIT 180

QY 263 ASSIFFTLMDMFTWEPKARLDKQGVNMTSGHNDOSOMLOYDLLVPTKVTGIIITOGA 322
 DB 181 ASSIFFTLMDMFTWEPKARLDKQGVNMTSGHNDOSOMLOYDLLVPTKVTGIIITOGA 240

QY 323 KDGCHQVFGSKTIAYSDGGEHMTVYODEKORKDKVPGCNPDNTHKKNYIDPPIYARIH 382
 DB 241 KDGCHQVFGSKTIAYSDGGEHMTVYODEKORKDKVPGCNPDNTHKKNYIDPPIYARIH 300

QY 383 RILPMSWYGRITLASSELLGCT 403
 DB 301 RILPMSWYGRITLASSELLGCT 321

RESULT 3
 US-09-237-981-21
 ; Sequence 21, Application US/09237981
 ; GENERAL INFORMATION:
 ; APPLICANT: Quettermous, Thomas
 ; APPLICANT: Hogan, Bridgid
 ; APPLICANT: Snodgrass, H. Ralph
 ; APPLICANT: Zupancic, Thomas J.
 ; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 ; TITLE OF INVENTION: CELL LOCUS-1
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/237,981
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/659,235
 ; FILING DATE: 05-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.
 REGISTRATION NUMBER: 28,462
 REFERENCE/DOCKET NUMBER: 8907-0034-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 Pennle
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 321 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-09-237-981-21

Query Match 75.8%; Score 1717; DB 16; Length 321;
 Best Local Similarity 98.4%; Pred. No. 5,6e-151;
 Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 83 CSGPLGIEGGIISNOGITASSSTRHALFGLQKWPYARLNKKGLINAMTAENDRMKWT 142
 DB 1 CSGPLGIEGGIISNOGITASSSTRHALFGLQKWPYARLNKKGLINAMTAENDRMKWT 60
 QY 143 QINLQRMKRYTGVITOGARRISPEYIKFYKIAVSDGKTWAMYKVGTDNEDVFRGNID 202
 DB 61 QINLQRMKRYTGVITOGARRISPEYIKFYKIAVSDGKTWAMYKVGTDNEDVFRGNID 120
 QY 203 NNTPYANSTPTPIKAQYVRLYPQVCRHCTLRMELLCGCELSGSEPLGMSGHIQDYQIT 262
 DB 121 NNTPYANSTPTPIKAQYVRLYPQVCRHCTLRMELLCGCELSGSEPLGMSGHIQDYQIT 180
 QY 263 ASSIFRLINDMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQVLDLVPVTGIIITOGA 322
 DB 181 ASSIFRLINDMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQVLDLVPVTGIIITOGA 240
 QY 323 KDFGVQVPGSYKLAYSNDGEHWYVODEKQKRDKNVGFNDHRRKNVIDPPIYARHI 382
 DB 241 KDXGHVQVPGSYKLAYSNDGEHWYVODEKQKRDKNVGFNDHRRKNVIDPPIYARHI 300
 QY 383 RILPMSWGRITLASLIGCT 403
 DB 301 RILPMSWGRITLASLIGCT 321

RESULT 4
 US-60-207-315-523
 ; Sequence 523, Application US/60207315
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
 ; FILE REFERENCE: CLO00601
 ; CURRENT APPLICATION NUMBER: US/60/207,315
 ; CURRENT FILING DATE: 2000-05-30
 ; NUMBER OF SEQ ID NOS: 528
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 523
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: HUMAN
 US-60-207-315-523

Query Match 62.1%; Score 1406.5; DB 23; Length 362;
 Best Local Similarity 88.8%; Pred. No. 4,5e-122;
 Matches 262; Conservative 3; Mismatches 7; Indels 23; Gaps 2;

QY 44 NINCEVPEPCCKNGGICITDLYANVSCCEGPEFGMRNCQYKSGPLGIEGGIISNOGITASS 103
 DB 52 DINECEVPEPCCKNGGICITDLYANVSCCEGPEFGMRNCQYKSGPLGIEGGIISNOGITASS 111

QY 104 THRALFGLQKWPYARLNKKGLINAMTAENDRMK-----RMIOINLQRMKRV 152
 DB 112 THRALFGLQKWPYARLNKKGLINAMTAENDRMKRWIOELRGCHLGWMOINLQRMKRV 171
 QY 153 TGVITOGARRISPEYIKFYKIAVSDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSTFT 212
 DB 172 TGVITOGARRISPEYIKFYKIAVSDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSTFT 231
 QY 213 PPIKAQYVRLYPQVCRHCTLRMELLCGCELSGSEPLGMSGHIQDYQITASSIFRTLNM 272
 DB 232 PPIKAQYVRLYPQVCRHCTLRMELLCGCELSGSEPLGMSGHIQDYQITASSIFRTLNM 291
 QY 273 DMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQVLDLVPVTGIIITOGA 315
 DB 292 DMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQVLDLVPVTGIIITOGA 346

RESULT 5
 US-60-230-435-1519
 ; Sequence 1519, Application US/60230435
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
 ; FILE REFERENCE: CLO00768
 ; CURRENT APPLICATION NUMBER: US/60/230,435
 ; CURRENT FILING DATE: 2000-09-06
 ; NUMBER OF SEQ ID NOS: 2991
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1519
 ; LENGTH: 311
 ; TYPE: PRT
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(311)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-60-230-435-1519

Query Match 62.1%; Score 1405.5; DB 23; Length 311;
 Best Local Similarity 89.1%; Pred. No. 4,6e-122;
 Matches 262; Conservative 2; Mismatches 7; Indels 23; Gaps 2;

QY 45 INCEVPEPCCKNGGICITDLYANVSCCEGPEFGMRNCQYKSGPLGIEGGIISNOGITASS 104
 DB 2 INCEVPEPCCKNGGICITDLYANVSCCEGPEFGMRNCQYKSGPLGIEGGIISNOGITASS 61
 QY 105 HRALFGLQKWPYARLNKKGLINAMTAENDRMK-----RMIOINLQRMKRV 153
 DB 62 HRALFGLQKWPYARLNKKGLINAMTAENDRMKRWIOELRGCHLGWMOINLQRMKRV 171
 QY 154 TGVITOGARRISPEYIKFYKIAVSDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSTFT 213
 DB 122 TGVITOGARRISPEYIKFYKIAVSDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSTFT 181
 QY 214 PPIKAQYVRLYPQVCRHCTLRMELLCGCELSGSEPLGMSGHIQDYQITASSIFRTLNM 273
 DB 182 PPIKAQYVRLYPQVCRHCTLRMELLCGCELSGSEPLGMSGHIQDYQITASSIFRTLNM 241
 QY 274 DMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQVLDLVPVTGIIITOGA 315
 DB 242 DMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQVLDLVPVTGIIITOGA 346

RESULT 6
 US-09-582-340-2
 ; Sequence 2, Application US/09582340
 ; GENERAL INFORMATION:
 ; APPLICANT: INSERM
 ; APPLICANT: INSTITUT CURIE

APPLICANT: CNRS
TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants
FILE REFERENCE: Lactadherin
CURRENT APPLICATION NUMBER: US/09/582,340
CURRENT FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 2
LENGTH: 387
TYPE: PRT
ORGANISM: Homo sapiens
US-09-582-340-2

Query Match 45.1%; Score 1020.5; DB 19; Length 387;
Best Local Similarity 48.8%; Pred. No. 3,4e-96;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;

QY 3 CTNPGCHNGTTC-EISEAYRGDFEIGVYCKPGRFGNGICOHNINECEVEPCNGICTD 61
DB 27 CSKNCHNGGLCEIEISQEVGRGVFESYCTCTLKGYAGNH----- 65
QY 62 LVANYSCECPGEFMRGRCQYKCSGLPIEGGIIISNOQITASTHRAFLGLOKWPYYARL 121
DB 66 -----CETKCYEPLEGMEGNINANSQIAASVRYFELQHWPELARL 108
QY 122 NKKGLINAMTAENDRKRMIQINLORKRMYGVITGAKRIGSPYIKFYKIASYNDGK 181
DB 109 NRAGVNAWTPSSND-NPMIOYNLLRRMYTVGTGASRLASHKAFVAVSLNGH 167
QY 182 TW-AWYKFKGTNEDEMYFRGNIDNTPYANSFPPIKAOYRYLYPOVCRHCTLRMELLC 240
DB 168 EPEFDHDKKKHEFY--GWNKNNAVHNLFEFVEAOYRLVTPSCHTACTLRMELLC 225
QY 241 EISGSEPLGKSGHIDYQITASSIFPTLMMDWTEPRKARLDKOGVNAWTSQNDQ 300
DB 226 ELNGCANPLGLKNNISIPDRQITASSSYKTWGLHFSNPNPVALDKOGVFNMAWAGSYGN 285
QY 301 SOMLOVDLVPYKTYGTTGAGFQGVGYSKYKLAWSNDGSHMTYVODEKORRDKVFO 360
DB 286 DQMLQVDDGSSKEVLTGIIITGASANGSYQVFAASYKVAWSNDSNMTVEYDPTGSSKIFP 345
QY 361 GNEBNDTRKKNVIDPPIYARHRIILPWSMYGRITLASELLGC 402
DB 346 GNMGNHSHKKNLFEPTILAYVRIPLVAMHRIALRELLGC 387

RESULT 7

US-08-480-229B-20
Sequence 20 Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, R. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B

FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: protein
MOLECULE TYPE: protein
US-08-480-229B-20

Query Match 41.9%; Score 948; DB 8; Length 320;
Best Local Similarity 54.7%; Pred. No. 1,4e-79;
Matches 175; Conservative 50; Mismatches 95; Indels 0; Gaps 0;

QY 83 CSGPLGIEGGIISNOQITASTHRAFLGLOKWPYYARLNKKGLINAMTAENDRKRMI 142
DB 1 CSTOLGMEGALINDSQTSSAYVMGFGLQRMGPBLARLYRTGIYVMMHASNVDXKPMI 60
QY 143 QINLORKRMYGVITGAKRIGSPYIKFYKIASYNDGKTAMRYKGTNEDMYFRGND 202
DB 61 QYNLLRRKMYSGVMTGASRAGRAEYIKTKFVAISLDGAKREFIODESGDKEFLGND 120
QY 203 NMTYANSFPPIKAOYRYLYPOVCRHCTLRMELLCGSCSPGLGMSGHIDYQIT 262
DB 121 NNSLKYVMEFTELAQIRLYPYVCHGCTLRMELLCGSCLEPGLKNNTPSSQMS 180
QY 263 ASSIFPTLMMDWTEPRKARLDKOGVNAWTSQNDQ SOMLOVDLVPYKTYGTTG 322
DB 181 ASSSKYTNLRAFGWYFHLGRLDNGKINMTAOSNAKEMLDLQGVQYVGLITGQA 240
QY 323 KDFGHVQVSYKLAWSNDGSHMTYVODEKORRDKVQGNPNDOTHRKNVIDPPIYAR 382
DB 241 RQFHIQIVESTIVANSDGQVMTYVXKEBQSSKYVQGNLNNSHKKNLFEKPMARY 300
QY 383 RILPWSMYGRITLASELLGC 402
DB 301 RVLPSVMAHRIILRELLGC 320

RESULT 8

US-08-659-235B-20
Sequence 20 Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, R. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:


```

Query Match          36.1%; Score 816.5; DB 23; Length 203;
Best Local Similarity 97.9%; Pred. No. 1.3e-67;
Matches 141; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPCPTNPGHNGTCEISEAYRSDTFIYVCKCPRGFNGIHCOHNIKECEVEPCKNIGICT 60
DB 54 GPCPTNPGHNGTCEISEAYRSDTFIYVCKCPRGFNGIHCOHNIKECEVEPCKNIGICT 113
QY 61 DLVANYSCGPCGPFMRNCOYKCSGPLGIEGGIISNOQITASSTRALFGLQKWPYYAR 120
DB 114 DLVANYSCGPCGPFMRNCOYKCSGPLGIEGGIISNOQITASSTRALFGLQKWPYYAR 173
QY 121 LNKKGILINAWTAENDRMRWIOI 144
DB 174 LNKKGILINAWTAENDRMRWIOI 196

RESULT 11
US-60-230-435-1806
; Sequence 1806, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; NUMBER OF SEQ. ID NOS.: 2991
; SOFTWARE: PASTSEQ for Windows Version 4.0
; SEQ. ID NO. 1806
; LENGTH: 185
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(185)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-230-435-1806

Query Match          35.9%; Score 813.5; DB 23; Length 185;
Best Local Similarity 98.6%; Pred. No. 2.2e-67;
Matches 141; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GPCPTNPGHNGTCEISEAYRSDTFIYVCKCPRGFNGIHCOHNIKECEVEPCKNIGICT 60
DB 44 GPCPTNPGHNGTCEISEAYRSDTFIYVCKCPRGFNGIHCOHNIKECEVEPCKNIGICT 103
QY 61 DLVANYSCGPCGPFMRNCOYKCSGPLGIEGGIISNOQITASSTRALFGLQKWPYYAR 120
DB 104 DLVANYSCGPCGPFMRNCOYKCSGPLGIEGGIISNOQITASSTRALFGLQKWPYYAR 163
QY 121 LNKKGILINAWTAENDRMRWIOI 143
DB 164 LNKKGILINAWTAENDRMRWIOI 185

RESULT 12
US-09-237-981-29
; Sequence 29, Application US/09237981
; GENERAL INFORMATION:
; APPLICANT: Quattermus, Thomas
; APPLICANT: Hogen, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCOS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas

```

```

CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ. ID NO.: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-09-237-981-29

Query Match          34.7%; Score 784.5; DB 16; Length 221;
Best Local Similarity 92.5%; Pred. No. 1.4e-64;
Matches 139; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 1 GPCPTNPGHNGTCEISEAYRSDTFIYVCKCPRGFNGIHCOHNIKECEVEPCKNIGICT 60
DB 76 GPCPTNPGHNGTCEISEAYRSDTFIYVCKCPRGFNGIHCOHNIKECEVEPCKNIGICT 135
QY 61 DLVANYSCGPCGPFMRNCOYKCSGPLGIEGGIISNOQITASSTRALFGLQKWPYYAR 120
DB 136 DLVANYSCGPCGPFMRNCOYKCSGPLGIEGGIISNOQITASSTRALFGLQKWPYYAR 195
QY 121 LNKKGILINAWTAENDRMRWIOI 146
DB 196 LNKKGILINAWTAENDRMRWIOI 220

RESULT 13
US-08-482-596A-3
; Sequence 3, Application US/08482596A
; GENERAL INFORMATION:
; APPLICANT: Ceriani, Roberto L.
; APPLICANT: Peterson, Jerry A.
; APPLICANT: Larocca, David J.
; TITLE OF INVENTION: FUSION PROTEIN WITH 46 KDALTON
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: Suite 301
; CITY: One Westlakes, Berwyn
; STATE: Valley Forge
; COUNTRY: Pennsylvania
; ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 3.5"
COMPUTER: IBM PC Compatible

```

US-08-15
US-08-1482-596A-4
Sequence 4, Application US/08482596A
GENERAL INFORMATION:
APPLICANT: Ceriani, Roberto L.
APPLICANT: Peterson, Jerry A.
APPLICANT: Latocca, David J.
TITLE OF INVENTION: PSION PROTEIN WITH 46 KDALTON
TITLE OF INVENTION: KIT & METHODS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rather & Prestia
STREET: Suite 301
CITY: One Westlakes, Berryn
STATE: Valley Forge
COUNTRY: Pennsylvania
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk 3.5"
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: PatentIn #1.0
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,596A
FILING DATE: June 7, 1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930

REFERENCE/DOCKET NUMBER: CRFC-046
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 407-0700
TELEFAX: (610) 407-0701
TELEX: N.A.
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE:
US-08-482-596A-4

Query Match 21.8%; Score 494; DB 8; Length 218;
Best Local Similarity 45.2%; Pred. No. 1.3e-37;
Matches 98; Conservative 34; Mismatches 63; Indels 2; Gaps 1;
QY 186 YKVGKTNEDMVRGNIIDNTPYANSFTPIKAQYRLIPQVCRHCTLRMELLGCELSGC 245
DB 1 FKGNSTRNMYFNGNSDASTIKENQFDPPIVARYIRISPTRAYNRPRLRLLEQCEVNGC 60
QY 246 SEPLGKSGHIDYQITFASISIFRTLNMDFTWEPKRLDKQGVNMTSGNDOSOWLO 305
DB 61 STPLGMEKGIENKQITFASISFKSMWGDY--WEPPFARLINOGRVNAWQAKANNKOWLE 118
QY 306 VDLIVPTKVTGLITOGAKDFGHVOFGYSYKLAYSNDGEHMTYODEKORKDKVFOGNFDM 365
DB 119 IDLLKIKITAITITOGCKLSSEMYVASYTHYSEGVEMKPYRLKSSMDKIFEGNTNT 178
QY 366 DTRKRVVIDPPIYARHRIPLMSYGRITLASELIGC 402
DB 179 KGHVKNFENPILISRFIRVLPKTNOSIALRLLELFC 215

Search completed: May 23, 2001, 08:36:45
Job time: 121 sec

Wed May 23 08:50:20 2001

us-09-237-981-14_copy_108_513.clo.rapn

Page 3

```

.PRIOR APPLICATION NUMBER: 09/649,167
} PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/668,680
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/695,618
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 09/728,711
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
PRIOR FILING DATE: 2000-03-14
NUMBER OF SEQ ID NOS: 172
SOFTWARE: Custom
SEQ ID NO: 74
LENGTH: 201
TYPE: PRT
ORGANISM: Homo sapiens
CCT-0501-10484-74

```

```

Query Match Score 166; DB 1; Length 201;
Best Local Similarity 33.6%; Pzed. No. 6.6e-09;
Matches 38; Conservative 24; Mismatches 43; Indels 8; Gaps 3

OY 279 PKRATLDQGVNMTSGHNDOSOMLDLVPKRCITITQACRFGHVCVQSGYKAY 338
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 56 PGRSTRNKRDAGAGMTPLVSNKTKYQLOLDLGERHEVTAVAIQG--TGSDDWITLPM 113
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 339 SNDGEHWTYODEKQRDKV--FQGNFDNDTHRKKNATIDPELYARHRIILPMSW 389
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 114 SDGGRMKYQY---RREESTIWGPGGTANDSVYHRIQPPERARTRRLPLAM 162
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

RESULT 7
 PCT-US01-01302-90
 Sequence 90, Application PC/TUS0101302
 GENERAL INFORMATION:
 APPLICANT: Human Genome Sciences, Inc., et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: P1707PCT
 CURRENT APPLICATION NUMBER: PCT/US01/01302
 CURRENT FILING DATE: 2001-01-17
 NUMBER OF SEQ ID NOS: 192
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 90
 LENGTH: 213
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (1)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (17)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (179)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (195)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (205)
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 PCT-US01-01302-90

```

Query Match      6.98; Score 156.5; DB 1, length 213;
Best Local Similarity 35.48; Pred. No. 5.7e-08;
Matches 28; Conservative 10; Mismatches 36; Indels 5; Gaps 1

OY 2 PCTPNCHNGTCTEISFARYKDTFIFGVCKCRGNGHICHOHNIINCEVEYPPCKNGTGITD 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
ob 98 PCSNNSCCSASACVATFTREI-----SFLCKCPGYSIGTETITGSGNNSDCHGTHIQ 152

```

```
QY      62 LVANYSCECPGEFMRNCQ  80
          | | : | | :
Db     153 DPIYPVICPAGYAGRCE 171
```

```

RESULT      8
PCT-US01-01321-1559
: Sequence 1559, Application PC/TUS0101321
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc., et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC0119C7
: CURRENT APPLICATION NUMBER: PCT/US01/01321
: NUMBER OF SEQ ID NOS: 2181
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1559
: LENGTH: 213
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (17)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (179)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (195)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (205)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01321-1559

```

Query Match	6.9%	Score 156.5	DB 1	Length 213
Best Local Similarity	35.4%	Pred. No. 5,7e-08		
Matches	28	Conservative	10	Mismatches 36
			Indels	5
			Gaps	1

OY	2	PCPCPNCHAGGICGCELSFAYRGDFICGKCGRGNGIHOHMINCFEPEPKNGICITD	61
Db	98	PCSNACQGSATCVNTPFEER-----SFLCKPPIISGICETITSGCKNSQMHGICHQ	152
OY	62	LVANISCECPCEGHEMRNQ	80
Db	153	DPILVPCICPAGYAGRFE	171

```

RESULT          9      379-4
US-09-808-689-4
: Sequence 4, Application US/09808689
: GENERAL INFORMATION:
: APPLICANT: Fong, Sherman
: APPLICANT: Audrey Goddard
: APPLICANT: Gunney, Austin L.
: APPLICANT: Tumas, Daniel
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
: TITLE OF INVENTION: RELATED DISEASES
: FILE REFERENCE: P1624RC1
: CURRENT APPLICATION NUMBER: US/09/808,689
: CURRENT FILING DATE: 2001-03-14
: Prior application data removed - consult PAIN or file wrapper
: NUMBER OF SEQ. ID NOS: 39
: SEQ. ID NO. 4
: LENGTH: 379
: TYPE: PRT
: ORGANISM: Homo sapiens

```

Wed May 23 08:50:20 2001

us-09-237-981-14_copy_108_513.clo.rapn

Page 4

US-09-808-689-4

Query Match 6.2%; Score 140.5; DB 5; Length 379;
Best Local Similarity 32.7%; Pred. No. 3.9e-06;
Matches 34; Conservative 7; Mismatches 30; Indels 33; Gaps 5;

QY 5 PNPCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHININECEVEPCKNGICTDLVA 64
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 183 PGCGRNGGFCNERR-----ICECPDGFHGPCHCKAL--C-TPRCMNGSLC---VT 226
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 65 NYSCECPGEFMGRNCOY-----KCSGPLGIEG 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 227 PGFCICPPGFYGVNCDKANCSTTCFNGTCTFYPGKCTICPPGLG 270
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 10

US-09-423-844-4

; Sequence 4, Application US/09423844
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Chen, Jian
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth
; APPLICANT: Pennica, Diane
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P1618P1E
; CURRENT APPLICATION NUMBER: US/09/423,844
; CURRENT FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-423-844-4

Query Match 6.2%; Score 140.5; DB 5; Length 379;
Best Local Similarity 32.7%; Pred. No. 3.9e-06;
Matches 34; Conservative 7; Mismatches 30; Indels 33; Gaps 5;

QY 5 PNPCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHININECEVEPCKNGICTDLVA 64
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 183 PGCGRNGGFCNERR-----ICECPDGFHGPCHCKAL--C-TPRCMNGSLC---VT 226
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 65 NYSCECPGEFMGRNCOY-----KCSGPLGIEG 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 227 PGFCICPPGFYGVNCDKANCSTTCFNGTCTFYPGKCTICPPGLG 270
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 11

US-09-403-296A-4

; Sequence 4, Application US/09403296A
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Audrey Goddard
; APPLICANT: Gurney, Austin L.
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
; TITLE OF INVENTION: RELATED DISEASES
; FILE REFERENCE: P1624R2E
; CURRENT APPLICATION NUMBER: US/09/403,296A
; CURRENT FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/100,858

; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/19437
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-403-296A-4

Query Match 6.2%; Score 140.5; DB 5; Length 379;
Best Local Similarity 32.7%; Pred. No. 3.9e-06;
Matches 34; Conservative 7; Mismatches 30; Indels 33; Gaps 5;

QY 5 PNPCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHININECEVEPCKNGICTDLVA 64
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 183 PGCGRNGGFCNERR-----ICECPDGFHGPCHCKAL--C-TPRCMNGSLC---VT 226
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 65 NYSCECPGEFMGRNCOY-----KCSGPLGIEG 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 227 PGFCICPPGFYGVNCDKANCSTTCFNGTCTFYPGKCTICPPGLG 270
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 12

US-09-380-139A-4

; Sequence 4, Application US/09380139A
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC. et al.
; APPLICANT: Chen, Jian
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth
; APPLICANT: Pennica, Diane
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: 10466-04
; CURRENT APPLICATION NUMBER: US/09/380,139A
; CURRENT FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US 98/19330
; PRIOR FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-380-139A-4

Query Match 6.2%; Score 140.5; DB 5; Length 379;
Best Local Similarity 32.7%; Pred. No. 3.9e-06;
Matches 34; Conservative 7; Mismatches 30; Indels 33; Gaps 5;

QY 5 PNPCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHININECEVEPCKNGICTDLVA 64
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 183 PGCGRNGGFCNERR-----ICECPDGFHGPCHCKAL--C-TPRCMNGSLC---VT 226
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 65 NYSCECPGEFMGRNCOY-----KCSGPLGIEG 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
DB 227 PGFCICPPGFYGVNCDKANCSTTCFNGTCTFYPGKCTICPPGLG 270
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 13

PCT-US01-01302-97

; Sequence 97, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17


```
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 97
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01302-97
```

```
Query Match          6.1%; Score 137; DB 1; Length 122;
Best Local Similarity 41.3%; Pred. No. 2e-06;
Matches 26; Conservative 8; Mismatches 23; Indels 6; Gaps 2;
```

```
QY 28 YVCKPGRFNGHICQHINECEVEPCCKNGIC---TDLVANYSCBCEPGEFNGRCQY--- 81
   : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 39 FECTCRGRGYGLRCEVSCVTCADPCFNGGLCVGADPDSAYICHPGFGSNCERVD 98
QY 82 KCS 84
   : ||
Db 99 RCS 101
```

```
RESULT 14
PCT-US01-01302-95
; Sequence 95, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 95
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (119)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (120)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (133)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (142)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-95
```

```
Query Match          5.9%; Score 133.5; DB 1; Length 156;
Best Local Similarity 35.7%; Pred. No. 5.8e-06;
Matches 25; Conservative 8; Mismatches 30; Indels 7; Gaps 2;
```

```
QY 3 CTNPNC-HNGGTCETISEAYRQDPTFYGVCCKPGRFNGHICQHINECEVEPCCKNGICTD 61
   : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 62 CAVNPLAHNG-----ECVADNISRGTGHCPCPPSOGDDCSVDVNECLSNPCPSTATCNM 115
QY 62 LVANYSCECP 71
   : ||
Db 116 TGGXXICKCP 125
```

```
RESULT 15
PCT-US01-01302-106
; Sequence 106, Application PC/TUS0101302
; GENERAL INFORMATION:
```

```
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 106
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01302-106
```

```
Query Match          5.5%; Score 123.5; DB 1; Length 201;
Best Local Similarity 28.7%; Pred. No. 7.1e-05;
Matches 29; Conservative 10; Mismatches 31; Indels 31; Gaps 5;
```

```
QY 3 CTNPNC-HNGGTCETISEAYRQDPTFYGVCCKPGRFNGHICQHINECEVEPCCKNGIC---C 59
   : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 63 CSAGLCFPHGRCVPGSA-----QPCHCPGPGFGPGRQYDVDECRTH---NGGCQHRG 111
QY 60 TDLVANYSCBCEPGEF-----MGR-NCQYKC 83
   : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 112 VNPBSYLCECKPGRFRLTDSRTCLAINSCALGNGCQHRG 132
```

```
Search completed: May 23, 2001, 08:36:54
Job time: 120 sec
```


Wed May 23 08:50:17 2001

us-09-237-981-14_copy_108_513.clo.ra1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:34:09 ; Search time 12.81 Seconds
(without alignments)
608,869 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Perfect score: 2264
Sequence: 1 GPCPTPNCNGSTCEISBAY.....MSWGRITLASELLGCTEER 406

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210657 residues

Total number of hits satisfying chosen parameters: 171963

Minimum DB seq length: 0
Maximum DB seq length: 406

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents: AA: *
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1717	75.8	321	2	US-08-480-229C-21
2	1717	75.8	321	2	US-08-659-235C-21
3	1020.5	45.1	387	2	US-08-162-402B-6
4	948	41.9	320	2	US-08-480-229C-20
5	948	41.9	320	2	US-08-659-235C-20
6	784.5	34.7	221	2	US-08-480-229C-29
7	784.5	34.7	221	2	US-08-659-235C-29
8	643	28.4	217	1	US-07-607-538C-3
9	643	28.4	217	1	US-07-607-538C-3
10	643	28.4	218	1	US-08-162-402B-3
11	643	28.4	218	1	US-07-607-538C-2
12	501	22.1	157	2	US-08-162-402B-13
13	498	22.0	159	2	US-08-162-402B-12
14	494	21.8	218	1	US-07-607-538C-4
15	494	21.8	218	1	US-08-162-402B-4
16	475.5	21.0	218	1	US-07-607-538C-5
17	475.5	21.0	218	1	US-08-162-402B-5
18	424	18.7	160	2	US-08-162-402B-11
19	414	18.3	160	2	US-08-162-402B-10
20	387	17.1	85	2	US-08-480-229C-1
21	387	17.1	85	2	US-08-659-235C-1
22	376	16.7	160	2	US-08-162-402B-14
23	368	16.3	109	1	US-08-111-939-23
24	344.5	15.2	160	2	US-08-162-402B-15
25	305.5	13.5	134	2	US-08-162-402B-17
26	304	13.4	110	1	US-08-111-939-24
27	293.5	13.0	159	2	US-08-162-402B-16

28	264	11.7	43	2	US-08-480-229C-24	Sequence 24, Appl
29	264	11.7	43	2	US-08-659-235C-24	Sequence 24, Appl
30	257	11.4	111	1	US-08-111-939-23	Sequence 23, Appl
31	252	11.1	85	2	US-08-480-229C-2	Sequence 2, Appl
32	252	11.1	85	2	US-08-659-235C-2	Sequence 2, Appl
33	232.5	10.3	109	1	US-08-111-939-21	Sequence 19, Appl
34	227	10.0	161	2	US-08-162-402B-19	Sequence 18, Appl
35	221.5	9.8	109	1	US-08-111-939-19	Sequence 18, Appl
36	220	9.7	156	2	US-08-162-402B-18	Sequence 18, Appl
37	216.5	9.6	109	1	US-08-111-939-20	Sequence 20, Appl
38	207.5	9.2	115	1	US-08-111-939-22	Sequence 22, Appl
39	198.5	8.8	157	2	US-08-872-855-6	Sequence 6, Appl
40	196.5	8.7	142	2	US-08-480-229C-25	Sequence 25, Appl
41	196.5	8.7	383	1	US-08-659-235C-25	Sequence 25, Appl
42	187	8.3	383	1	US-08-597-545-2	Sequence 2, Appl
43	187	8.3	385	1	US-08-457-135-2	Sequence 1, Appl
44	179	7.9	385	1	US-08-597-545-1	Sequence 1, Appl
45	179	7.9	385	1	US-08-457-135-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-480-229C-21
Sequence 21, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-229C-21
Query Match 75.8%; Score 1717; DB 2; Length 321;
Best Local Similarity 98.4%; Pred. No. 6,9e-135;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 83 CSSEPLGEGGIIISNOOITASTSTRALFGLQKWPYARLKKGLINAAENDKRWI 142
DB 1 CSSEPLGEGGIIISNOOITASTSTRALFGLQKWPYARLKKGLINAAENDKRWI 60
OY 143 QINLQKRWKVTGVTGQAKRIGSEPEIKFYKTAISNDCKTAMAKYKVGNEQWVRGND 202
DB 61 QINLQKRWKVTGVTGQAKRIGSEPEIKFYKTAISNDCKTAMAKYKVGNEQWVRGND 120
OY 203 NNTPYANSTPPPIKAQVRLYPQVCRHCTLMELLGCELSGSEPLGKMSGHIDYOIT 262
DB 121 NNTPYANSTPPPIKAQVRLYPQVCRHCTLMELLGCELSGSEPLGKMSGHIDYOIT 180
OY 263 ASSIFRTLMMDFTWEPKARLDKQKRVNMTSGHNDOSOMLOVLLVPTKYTGIIITOGA 322
DB 181 ASSIFRTLMMDFTWEPKARLDKQKRVNMTSGHNDOSOMLOVLLVPTKYTGIIITOGA 240
OY 323 KDFGHVQFVGSYKLAYSNDGEHMTVYQDEKQKRDVFOGNEFNDTHRRKNVIDPPIYARHI 382
DB 241 KDFGHVQFVGSYKLAYSNDGEHMTVYQDEKQKRDVFOGNEFNDTHRRKNVIDPPIYARHI 300
OY 383 RILPMSWGRTITLASELLGCT 403
DB 301 RILPMSWGRTITLASELLGCT 321

RESULT 2

US-08-659-235C-21
Sequence 21, Application US/08659235C
Patent No 5877281
GENERAL INFORMATION:
APPLICANT: Quartermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zuercher, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-659-235C-21

Query Match 75.8%; Score 1717; DB 2; Length 321;
Best Local Similarity 98.4%; Pred. No. 6,9e-135;

Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0:
OY 83 CSSEPLGEGGIIISNOOITASTSTRALFGLQKWPYARLKKGLINAAENDKRWI 142
DB 1 CSSEPLGEGGIIISNOOITASTSTRALFGLQKWPYARLKKGLINAAENDKRWI 60
OY 143 QINLQKRWKVTGVTGQAKRIGSEPEIKFYKTAISNDCKTAMAKYKVGNEQWVRGND 202
DB 61 QINLQKRWKVTGVTGQAKRIGSEPEIKFYKTAISNDCKTAMAKYKVGNEQWVRGND 120
OY 203 NNTPYANSTPPPIKAQVRLYPQVCRHCTLMELLGCELSGSEPLGKMSGHIDYOIT 262
DB 121 NNTPYANSTPPPIKAQVRLYPQVCRHCTLMELLGCELSGSEPLGKMSGHIDYOIT 180
OY 263 ASSIFRTLMMDFTWEPKARLDKQKRVNMTSGHNDOSOMLOVLLVPTKYTGIIITOGA 322
DB 181 ASSIFRTLMMDFTWEPKARLDKQKRVNMTSGHNDOSOMLOVLLVPTKYTGIIITOGA 240
OY 323 KDFGHVQFVGSYKLAYSNDGEHMTVYQDEKQKRDVFOGNEFNDTHRRKNVIDPPIYARHI 382
DB 241 KDFGHVQFVGSYKLAYSNDGEHMTVYQDEKQKRDVFOGNEFNDTHRRKNVIDPPIYARHI 300
OY 383 RILPMSWGRTITLASELLGCT 403
DB 301 RILPMSWGRTITLASELLGCT 321

RESULT 3

US-08-162-402B-6
Sequence 6, Application US/08162402B
Patent No 597237
GENERAL INFORMATION:
APPLICANT: CERTANT, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
APPLICANT: LARROCA, DAVID J.
TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder & Poplawski
STREET: 444 South Flower St., 19th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-6

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-659-235C-20

Query Match 41.9%; Score 948; DB 2; Length 320;
Best Local Similarity 54.7%; Pred. No. 3,7e-71;
Matches 175; Conservative 50; Mismatches 95; Indels 0; Gaps 0;

Oy 83 CSPLGTEGGITSMQOITASTSTRALFGLQKMYVYARLKKGLINAMTAENDRMKRM 142
Db 1 CSTQLMEGGAIDSOISASVYVGMGLQRMGPELRLRYGIYVAMHASNVDXSKPMI 60
Oy 143 QIMLRKMYVGVITOGAKRIGSPYKFKYKIVASNDCKTAMKYKGTNDWYFRGND 202
Db 61 QVNLKRMVSGVMTGASRAKREIKTFVATSLDGKXKEEFIDSGGDKREFLGND 120
Oy 203 NNPVANSFTPIKAQYVRLYPQVCRHCTLMELLCSELGSEPLGMSGHIOYQT 262
Db 121 NNSLKVMNPTLEQYIRLYPVSCHRCITREFELLCSELHGCLEPLDKNTIPDSQMS 180
Oy 263 ASSIFPTLMDFTWEPKRAKLDKQKNAVTSQNDOSOMLOVLLVPTVGSITTOGA 322
Db 181 ASSYKTMVLRAFGYPHLGRLDNQGKINAMTAQNSAKEMLOVLDGTOQVGTITIOGA 240
Oy 323 KDFGHVQVGSYKLYVANDGEHMTYVODEKORKDKVQGNFMDTHRKNVIDPITVARI 382
Db 241 RDGHIQVYVESYKVAHSDGVQWTYVXXEEOGSSKVFQGLDNHSHKKNIFKPFMARVY 300
Oy 383 RILPMGMYGRITLASELLGC 402
Db 301 RVLPSMNRITLRELLEGC 320

RESULT 6
US-08-480-229C-29
Sequence 29, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-480-229C-29

Query Match 34.7%; Score 784.5; DB 2; Length 221;
Best Local Similarity 92.5%; Pred. No. 8.3e-58;
Matches 135; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Oy 1 GPCTPNPGHNGCCEISEAYRDPFGYCKDPGPNFGHCOHINCEVEPCNGICT 60
Db 76 GCCTNPGHNGCCEISEAYRDPFGYCKDPGPNFGHCOHINCEVEPCNGICT 135
Oy 61 DLYAVYSCGCGEPMKNCQYKCSGPIGSGIISMQITASTSTRALFGLQKMYVYAR 120
Db 136 DLYAVYSCGCGEPMKNCQYKCSGPIGSGIISMQITASTSTRALFGLQKMYVYAR 195
Oy 121 LKKGLINAMTAENDRMKRM 146
Db 196 LKKGLINAMTAENDRMKRM 220

RESULT 7
US-08-659-235C-29
Sequence 29, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

FRAGMENT TYPE: internal
US-08-659-235C-29

Query Match 34.7%; Score 784.5; DB 2; Length 221;
Best Local Similarity 92.5%; Pred. No. 8.3e-58;
Matches 135; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
QY 1 GPCPTNPGHNGCGCEISAEKDTFTIGYCKPCRGRTGHCCHNTNCEGEVCKGCT 60
DB 76 GPCPTNPGHNGCGCEISAEKDTFTIGYCKPCRGRTGHCCHNTNCEGEVCKGCT 135
QY 61 DLVANSCEGCEGEMGNCQYKSGPLDIEGIIISNOQTASTTRALFLOKAYPYAR 120
DB 136 DLVANSCEGCEGEMGNCQYKSGPLDIEGIIISNOQTASTTRALFLOKAYPYAR 195
QY 121 LNKGLINANTAAENDMKRWICQINL 146
DB 196 LNKGLINANTAAENDMKRWICQINL 220

RESULT 8
US-07-607-538C-3
Sequence 3, Application US/07607538C
Patent No. 545031
GENERAL INFORMATION:
APPLICANT: Certant Dr., Roberto L.
APPLICANT: Peterson Dr., Jerry A.
APPLICANT: PETERSON, JERRY A.
TITLE OF INVENTION: DIFFERENTIATION ANTIGEN BINDING SPECIFICITY AND CLOTTING
TITLE OF INVENTION: POLYPEPTIDE WITH 46
TITLE OF INVENTION: FACTORS V AND VIII LIGHT-CHAIN HOMOLOGUES
TITLE OF INVENTION: FUSION PROTEIN, POLYPEPTIDE AND POLYBIO-
TITLE OF INVENTION: NUCLEOTIDE ENCODING THE POLYPEPTIDE, ANTI-
TITLE OF INVENTION: POLYPEPTIDE ANTIBODIES, KITS AND METHODS OF
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: V. Amzel & Assoc.
STREET: 2055 No. 5455031th Broadway
CITY: HAIOUT Creek
STATE: California
COUNTRY: USA
ZIP: 94596
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/607,538C
FILING DATE: 01-NOV-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: CRCC-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 943-1189
TELEFAX: (510) 943-1189
TELEX: N.A.
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE:
US-07-607-538C-3

Query Match 28.4%; Score 643; DB 1; Length 217;

Best Local Similarity 56.8%; Pred. No. 4.3e-46;
Matches 117; Conservative 30; Mismatches 59; Indels 0; Gaps 0;
QY 197 FRGNDNNTPVANSFTPPKQVRLYKPOVCRHCTLRMELLCGCELSGCEPLGMSGHI 256
DB 12 FGNMKNKNAHVNLFEPLVEADYVRLYPTSCFACFLRELLGCEELNCGANPLGLKNNST 71
QY 257 OYVOTASSIFRTLMMDMFTWEPKRAKLDKQGVNATSGHNDOSOMLOVDLAVPTKVTG 316
DB 72 PKQITASSYKTKWGLHFSWNPYSARLDKGNFNNAVAGSYGNDOMLOVDLSSSEKVEYG 131
QY 317 IITGAKDFGVHCVFVSGYKLAAYSNDGEHWYODEKOROKDFGQNFNDTHKRVNIDPP 376
DB 132 IITGARNFGSVGVAVSKVAASNDASWMTYODPRTSSKIFPGMDNHSKKNLFEPT 191
QY 377 IYARHRIILPMSWYGRITTLASELLCG 402
DB 192 IYARVRIILPVAHNRILALRELLCG 217

RESULT 9
US-08-162-402B-3
Sequence 3, Application US/08162402B
Patent No. 5972337
GENERAL INFORMATION:
APPLICANT: CERTANI, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
APPLICANT: LAROCCA, DAVID J.
TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
TITLE OF INVENTION: GLOBULE (HMF) ANTIGEN, FRAGMENTS & FUSION PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Priety, Schroeder & Poplawski
STREET: 444 South Flower St., 19th floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-3

Query Match 28.4%; Score 643; DB 2; Length 217;
Best Local Similarity 56.8%; Pred. No. 4.3e-46;
Matches 117; Conservative 30; Mismatches 59; Indels 0; Gaps 0;
QY 197 FRGNDNNTPVANSFTPPKQVRLYKPOVCRHCTLRMELLCGCELSGCEPLGMSGHI 256

Db 12 FVGNMKNKAAVHNLFEPTVEAOYVRLPTSCHTACTLRFELLGCELGNCANPLGKNNSI 71
QY 257 ODVOITASSIFRILNDMTWEPKARLDKOGKVNAMTSGHNDOSOMLQVDLLVPTKVTG 316
Db 72 PKOITASSSYKTWGLHLEFSMNPBYARLDKOGNFNMAVAGSYGNDOMLQVDLGSSKEVYG 131
QY 317 IITOGAKDFGHVDFVGSYKLAAYSDNGEHWTVODEKORKDKVFOGNEFNDTTRKNVTDPP 376
Db 133 IITOGARNFCSYGVFVASYKVAAYSDNSAMTVEYODPRTGSSKIFPGNMDHSHKKNLFPETP 191
QY 377 IVARHIRILPMWYGRITLASLLOC 402
Db 192 ILARYRILPVAMHNRILALRELLGC 217

RESULT 10
US-07-607-538C-2
; Sequence 2, Application US/07607538C
; Patent No. 5455031
; GENERAL INFORMATION:
; APPLICANT: CERTANI Dr., Roberto L.
; APPLICANT: PETERSON Dr., Jerry A.
; APPLICANT: LAROCCA, DAVID J.
; TITLE OF INVENTION: POLYPEPTIDE WITH 46
; TITLE OF INVENTION: DIFFERENTIATION ANTIGEN BINDING SPECIFICITY AND CLOTTING
; TITLE OF INVENTION: FACTORS V AND VIII LIGHT-CHAIN HOMOLOGUES, POLYRIBO-
; TITLE OF INVENTION: FUSION PROTEIN, POLYPEPTIDE AND POLYRIBO-
; TITLE OF INVENTION: NOCTURNE ENCODING THE POLYPEPTIDE, ANT
; TITLE OF INVENTION: USE THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: V Amzel & Assoc.
; STREET: 2055 No. 5455031th Broadway
; CITY: Walnut Creek
; STATE: California
; COUNTRY: USA
; ZIP: 94596
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/607/538C
; FILING DATE: 01-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: VIVIANA Amzel
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: CRFC-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 943-1931
; TELEFAX: (510) 943-1189
; TELEX: N.A.
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE:
US-07-607-538C-2

Query Match 28.4%; Score 643; DB 1; Length 218;
Best Local Similarity 56.8%; Pred. No. 4.3e-46;
Matches 117; Conservative 30; Mismatches 59; Indels 0; Gaps 0;
QY 197 FRCNDNNTPYANSFPPIKAGYVRLPYOVCRRHCTLMRLGCELGSCSEPLGKNSCHI 256

Db 13 FVGNMKNKAAVHNLFEPTVEAOYVRLPTSCHTACTLRFELLGCELGNCANPLGKNNSI 72
QY 257 ODVOITASSIFRILNDMTWEPKARLDKOGKVNAMTSGHNDOSOMLQVDLLVPTKVTG 316
Db 73 PKOITASSSYKTWGLHLEFSMNPBYARLDKOGNFNMAVAGSYGNDOMLQVDLGSSKEVYG 132
QY 317 IITOGAKDFGHVDFVGSYKLAAYSDNGEHWTVODEKORKDKVFOGNEFNDTTRKNVTDPP 376
Db 133 IITOGARNFCSYGVFVASYKVAAYSDNSAMTVEYODPRTGSSKIFPGNMDHSHKKNLFPETP 192
QY 377 IVARHIRILPMWYGRITLASLLOC 402
Db 193 ILARYRILPVAMHNRILALRELLGC 218

RESULT 11
US-08-162-402B-2
; Sequence 2, Application US/08162402B
; Patent No. 5972337
; GENERAL INFORMATION:
; APPLICANT: CERTANI, ROBERTO L.
; APPLICANT: PETERSON, JERRY A.
; APPLICANT: LAROCCA, DAVID J.
; TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
; TITLE OF INVENTION: GLOBULE (HMEG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pelly, Schroeder & Poplawski
; STREET: 444 South Flower St., 15th Floor
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: Fastseq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162/402B
; FILING DATE: 03-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Amzel, Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P66 38215
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-622-7700
; TELEFAX: 213-489-4210
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-162-402B-2

Query Match 28.4%; Score 643; DB 2; Length 218;
Best Local Similarity 56.8%; Pred. No. 4.3e-46;
Matches 117; Conservative 30; Mismatches 59; Indels 0; Gaps 0;
QY 197 FRCNDNNTPYANSFPPIKAGYVRLPYOVCRRHCTLMRLGCELGSCSEPLGKNSCHI 256
Db 13 FVGNMKNKAAVHNLFEPTVEAOYVRLPTSCHTACTLRFELLGCELGNCANPLGKNNSI 72
QY 257 ODVOITASSIFRILNDMTWEPKARLDKOGKVNAMTSGHNDOSOMLQVDLLVPTKVTG 316

Page 7

QY 364 DNDTHRKNVIDPPIYARHIRILPMSWYGRITLASELLG
||::||::: : ||::||::: ||::||::: ||::||::: ||::||:::

; Sequence 4, Application
; Patent NO. 5455031

Page 7

```

? APPLICANT: Ceriani Dr., Roberto L.
? APPLICANT: Peterson Dr., Jerry A.
? APPLICANT: Larocca, David J.
? TITLE OF INVENTION: POLYPEPTIDE WITH 46
? TITLE OF INVENTION: DIFFERENTIATION ANTIGEN BINDING SPECIFICITY AND CLOTTING
? TITLE OF INVENTION: FACTORS V AND VIII LIGHT-CHAIN HOMOLOGUES.
? TITLE OF INVENTION: FUSION PROTEIN, POLYNUCLEOTIDE AND POLYPEPTIDE
? TITLE OF INVENTION: NUCLEOTIDE ENCODING THE POLYPEPTIDE, ANTI-
? TITLE OF INVENTION: POLYPEPTIDE ANTIBODIES, KITS AND METHODS OF
? NUMBER OF SEQUENCES: 5
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: V. Amzel & Assoc.
? STREET: 2055 No. 545031th Broadway
? CITY: Walnut Creek
? STATE: California
? COUNTRY: USA
? ZIP: 94596
? COMPUTER READABLE FORM:
? MEDIUM TYPE: floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/607,538C
? FILING DATE: 01-NOV-1990
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Viviana Amzel
? REGISTRATION NUMBER: 30,930
? REFERENCE/DOCKET NUMBER: CRCC-004
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (510) 943-1931
? TELEFAX: (510) 943-1189
? TELEX: N.A.
? INFORMATION FOR SEQ. ID NO.: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 218 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? FRAGMENT TYPE:
? US-07-607-538C-4

Query Match 21.8%; Score 494; DB 1; Length 218;
Best Local Similarity 45.2%; Pred. No. 9.6e-34;
Matches 98; Conservative 34; Mismatches 83; Indels 2; Gaps 1;

QY 186 YKVGKTNEDMVFRCGNDNTNTPYANSFTPIKAQVRLYPQVCRRCCTLRMLLGCETLSCG 245
DB 1 FKGNSTRNVMYFNGNSDASTIKENQFDPPIVARIKISPTRAVNRPTRLRLDGCETVNGC 60
QY 246 SEPLGMSKSGHIDQYQITLASSIFRTLNDMFTWEPKARLDKQKYNAMTSGHNDOSQWLO 305
DB 61 STPLGMSKSGHIDQYQITLASSIFRTLNDMFTWEPKARLDKQKYNAMTSGHNDOSQWLO 118
QY 306 VDLVPTKVTGIIITQAGDFGHVQFVGSYKLAYSNDGEHMTYVODEKQKDKVFOGNDN 365
DB 119 IDLKIKKIKTALITQGGKSLSEMYKSYTYIHYSQGVEMKPYRLKSSMVDKIFEGNTNT 178
QY 366 DTHRKNVIDPPIVARIHRIPLMSWYGRITLASELLGC 402
DB 179 KGHVKNFNPILSRIRIVIPKTNQSGALRLLELFGC 215

```

```

? APPLICANT: PETERSON, JERRY A.
? APPLICANT: LARocca, DAVID J.
? TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
? TITLE OF INVENTION: GLOBULE (HMG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
? NUMBER OF SEQUENCES: 29
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pretty, Schroeder & Poplawski
? STREET: 444 South Flower St., 19th Floor
? CITY: Los Angeles
? STATE: CA
? COUNTRY: USA
? ZIP: 90071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/162,402B
? FILING DATE: 03-DEC-1993
? CLASSIFICATION: 435
? PRIOR APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Amzel, Viviana
? REGISTRATION NUMBER: 30,930
? REFERENCE/DOCKET NUMBER: P66 38215
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 213-622-7700
? TELEFAX: 213-489-4210
? TELEX:
? INFORMATION FOR SEQ. ID NO.: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 218 amino acids
? TYPE: amino acid
? STRANDEDNESS: unknown
? TOPOLOGY: unknown
? MOLECULE TYPE: peptide
? US-08-162-402B-4

Query Match 21.8%; Score 494; DB 2; Length 218;
Best Local Similarity 45.2%; Pred. No. 9.6e-34;
Matches 98; Conservative 34; Mismatches 83; Indels 2; Gaps 1;

QY 186 YKVGKTNEDMVFRCGNDNTNTPYANSFTPIKAQVRLYPQVCRRCCTLRMLLGCETLSCG 245
DB 1 FKGNSTRNVMYFNGNSDASTIKENQFDPPIVARIKISPTRAVNRPTRLRLDGCETVNGC 60
QY 246 SEPLGMSKSGHIDQYQITLASSIFRTLNDMFTWEPKARLDKQKYNAMTSGHNDOSQWLO 305
DB 61 STPLGMSKSGHIDQYQITLASSIFRTLNDMFTWEPKARLDKQKYNAMTSGHNDOSQWLO 118
QY 306 VDLVPTKVTGIIITQAGDFGHVQFVGSYKLAYSNDGEHMTYVODEKQKDKVFOGNDN 365
DB 119 IDLKIKKIKTALITQGGKSLSEMYKSYTYIHYSQGVEMKPYRLKSSMVDKIFEGNTNT 178
QY 366 DTHRKNVIDPPIVARIHRIPLMSWYGRITLASELLGC 402
DB 179 KGHVKNFNPILSRIRIVIPKTNQSGALRLLELFGC 215

Search completed: May 23, 2001, 08:35:33
Job time: 84 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on:

May 23, 2001, 08:36:59 ; Search time 20.7 seconds
(without alignments)
1596,895 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513

Perfect score: 2675

Sequence:

Scoring table:

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 481

Post-processing:

Database:

PIR 67:**
1: p1r1:**
2: p1r2:**
3: p1r3:**
4: p1r4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1173	43.9	409	2	PP47 protein - pig
2	1156	43.2	463	2	milk fat globule m
3	1154	43.1	427	2	ags protein precu
4	1145.5	42.8	427	2	PAS-6/7 protein pr
5	1119	41.8	401	2	glycoprotein antiq
6	643	24.0	218	2	milk fat globule p
7	464.5	17.4	216	2	factor Vllr-associ
8	268	10.0	473	2	adhesive plaque pr
9	253.5	9.5	308	3	epidermal growth f
10	252.5	9.4	293	2	neurogenic repetit
11	245	9.2	383	2	delta-like homeot
12	244.5	9.1	385	2	homoeotic protein d
13	244.5	9.1	385	2	predipocyte facto
14	239	8.9	387	2	Motch A protein -
15	204.5	7.6	252	2	hypothetical prote
16	182	6.8	379	2	mnt inhibitory fac
17	172	6.4	378	2	hypothetical prote
18	169.5	6.3	102	2	chondroitin sulfat
19	168.5	6.3	372	2	hypothetical prote
20	158.5	5.9	379	2	Abx-1 protein homo
21	148.5	5.6	407	1	coagulation factor
22	148	5.5	475	1	coagulation factor
23	147	5.5	259	2	hypothetical prote
24	135.5	5.1	429	2	hypothetical prote
25	135	5.0	443	2	coagulation factor
26	134.5	5.0	149	1	discoaldin I chain
27	134.5	5.0	149	1	discoaldin I chain
28	134.5	5.0	253	1	discoaldin I chain
29	131.5	4.9	356	2	hypothetical prote

30	131	4.9	470	2	A40697	63K sperm flagella
31	131	4.9	477	2	US0597	t-plasminogen acti
32	129.5	4.8	253	1	DLD01A	discoaldin I chain
33	129.5	4.8	264	2	T16271	hypothetical prote
34	126.5	4.7	396	1	KXB02	plasma protein Z -
35	125.5	4.7	409	2	T29517	hypothetical prote
36	123	4.6	356	2	A25918	thrombomodulin - b
37	123	4.6	422	1	KFH02	plasma protein 2 p
38	123	4.6	461	1	KFH02	coagulation factor
39	122.5	4.6	251	2	A55035	cysteine-rich prot
40	121	4.5	447	2	A59321	mucin - rat (fragm
41	116	4.3	293	2	T22919	hypothetical prote
42	116	4.3	402	2	S42367	Iag-2 protein - Ca
43	115	4.3	293	2	T09065	hypothetical prote
44	115	4.3	459	2	J00419	coagulation factor
45	114	4.3	416	1	KFB0	coagulation factor

ALIGNMENTS

RESULT 1
T11743
PP47 protein - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C/Accession: T11743
R:Brustlin, M.; Vogel, T.; Calveste, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toe
Biol. Reprod. 58, 1057-1064, 1998
A>Title: Molecular cloning and characterization of P47, a novel boar sperm-associated
A:Accession: T11743
A:Release: 1999-07-21/1925; M01D:98206817
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-409 <ENS>
A:Cross-references: EMBL:T11683; NID:g2652927; PIDN:CAA72379.1; PID:g2652928
A:Experimental source: testis
A:Function:
A:Description: may be involved in membrane remodeling and/or function as a zona pellu
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:6-40/Domain: EGF homology <EGF>

Query Match	43.9%	Score 1173;	DB 2;	Length 409;
Best local similarity	48.0%	Pred. No. 8.1e-79;		
Matches 223;	Conservative	63;	Mismatches 111;	Indels 68;
			Gaps 8;	
QY 23	GDICDPNCPNCEGICL-----PCLAVGSFSCPCPDGPTDPCNSVVEVASDEEPTSA	75		
DB 3	GDICDSSICLNGGICLLDDDPQK-----FHCICPEGPTGLICNE-----TEK	45		
QY 76	GPCTPNCPNCGNCE--ISEAYRGDFIGYCKCPRGFNGIHCHONINCEVEPCCKNGIC	134		
DB 46	GPCFNPCHNDACEVYIDAHRGDVFTEYICKCPHGYSIH-----	86		
QY 135	TDLVAVNSCECPDEFNRCYCKSGPLGIEGIIISNOQITFASSTHRLFGLOKWPYYA	194		
DB 87	-----CETICNAPIGMETGAIADFOISSSMHGFGIORNAPELA	127		
QY 195	RLNKGLINMTAENDRWKRWICINLQRCRWYGVITQGAARIGSEYKFKIAYNSD	254		
DB 128	RLHRAGIYNMTASNDR--NPVIGVNLRRMRVYGVITGASRAGSAEYKFKIAYSTD	186		
QY 255	GRTWAMKYVVGTE--DAVEFGNIDNMTPYANSFPTPIAQVRLAYOVQVRRKRLREL	312		
DB 187	GKRFQF--IGAEESGDKLFMGNDNSGLVNLFEVPLEVQVYVLAIVIIICRCLREL	244		
QY 313	LCCEISGCSPELQKSGHIOYQITASSIFRTLNMDMTWEPKRRARLDKQGVNMTSGH	372		
DB 245	LCCEISGCAEPLKDKNTIPNKQITASSFYRTGSAFVAFYARLDNGKRFNMTQOS	304		
QY 373	NDOSQMLQVLDLVPKRVTSIITGAKQDFGVQSVSKIAYNSDGEHWYVYDDEKORDK	432		

Db 305 NSASEMIQIDLGSRVRYGITQCARDFGHIOYVAAYVAVSDGVSWTEYRDQALEGK 364

OY 433 VQGFNFNDTRKRVNIDPPYARIRIRILPMSVWGRTITLASELLGC 477

Db 365 IFPGNLNNSHKKRMETPELREVRILPVAHNRITLRELLGC 409

RESULT 2

A36479

A:Accession: A36479

A:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

R:Stubbs, J.D.; Lekutis, C.; Singer, K.L.; Bul, A.; Yuzuki, D.; Srinivasan, U.; Parry, G

A:Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the ext

A:Reference number: A36479; MUID:91046008

A:Accession: A36479

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-463 <STU>

A:Cross-references: GB:M38337; NID:9199142; PIDN:AAA9534.1; PID:9199143

C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology: EGF homolo

F:28-60/Ds: membrane protein <EG1>

F:28-60/Ds: EGF homology <EG2>

F:68-107/Ds: EGF homology <EG3>

F:147-303/Ds: discoidin I amino-terminal homology <DN1>

F:307-463/Ds: discoidin I amino-terminal homology <DN2>

Query Match

Best Local Similarity 43.28; Score 1156; DB 1; Length 463;

Matches 222; Conservative 72; Mismatches 142; Indels 22; Gaps 8;

OY 23 GDICPNPCNGGICLPGIAGVSGSCPCPGFTDPCSVYEVASDEEPTSAAGCPNP 82

Db 25 GDFCSSLCLNGGICLTG- QNDIYCLCPGFTGLVGN- -----TERGCSNP 72

OY 83 CHNGTCEIS- EAYRGDTFEGYCKPFGNGHCHNHINECEVEPCXKNGICITDLVANY 141

Db 73 CYNDAKCLVETDQRDIFTEYICQCPVSGSGHCEETNYND- -GEYFTAVPNT 129

OY 142 SCECPGEF- MGRNCOYACGPIGIEGGLISNQITASTTRALFGLQKWPYARLNKK 199

Db 130 AVPPAPPTPLSLNMLASRSTOLMEGALADSQISASVYMGFGLQRMGPGLARLYRT 189

OY 200 GLINAMTAANDMKRMWIOINLQKRMVGVITQAKRIGSPYIKFYIAYSDNGRTWA 259

Db 190 GIVNMAHMSYDS- KFWIOVNLKRMVSGVMTGASRAGAEYLTFFVAASLDGRFE 248

OY 260 MYKVGINEMVFRGNIDNNTPYANSTPPIKAQVRLYPQVCRHRTLMELGCELSG 319

Db 249 FTQDE- SGGRKFLGNDLNSLKVMNPTLEAOYIRLYPVSCRHCTLRELLGCELSG 307

OY 320 CSEPLGKSGHIDVOITASSIFRTLMDFEWRKARLDKQKVNANTSGHDSQML 379

Db 308 CLEPGLKNNITPDSQMSASSYKTMWLFQWYHILGRDLNOKINAMTAQSNKAEWL 367

OY 380 QVDLLVTKVYITITQAKDGHVQFVSGKLAISNDEHWTYODEKORAKYFQGNED 439

Db 368 QVDLGTORQVYITITQAKDGHVQFVSGKLAISNDEHWTYODEKORAKYFQGNED 425

OY 440 NTHKRVNIDPPYARIRIRILPMSVWGRTITLASELLGC 477

Db 426 NNSHKKNIPEKPMARYAVLPVSMNRITLRELLGC 463

RESULT 3

JC4915

A: protein precursor - rat

N: Alternate names: O-acetyl-GD3 ganglioside

C: Species: Rattus norvegicus (Norway rat)

C: Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 20-Jun-2000

C:Accession: JC4915

R:Oura, K.; Nara, K.; Watanabe, Y.; Kohn, K.; Tai, T.; Sanai, Y.

Biochem. Biophys. Res. Commun. 225, 932-938, 1996

A:Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.

A:Reference number: JC4915; MUID:96374422

A:Accession: JC4915

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-427 <OCU>

A:Cross-references: DDBJ:D84068; NID:91620006; PIDN:BAH12210.1; PID:91620007

A:Experimental source: CST cell

C:Comment: This protein is required for the O-acetylation of disialoganglioside siall

C:Genes:

A:Superfamily: milk fat globule protein; discoidin I amino-terminal homology: EGF hom

F:1-21/Ds: signal sequence <status predicted> <SIG>

F:28-60/Ds: EGF homology <EG1>

F:68-107/Ds: EGF homology <EG2>

F:110-267/Ds: discoidin I amino-terminal homology <DN1>

F:271-427/Ds: discoidin I amino-terminal homology <DN2>

Query Match

Best Local Similarity 43.18; Score 1154; DB 2; Length 427;

Matches 221; Conservative 62; Mismatches 119; Indels 54; Gaps 7;

OY 23 GDICPNPCNGGICLPGIAGVSGSCPCPGFTDPCSVYEVASDEEPTSAAGCPNP 82

Db 25 GDFCSSLCLNGGICLTG- QNDIYCLCPGFTGLVGN- -----TERGCSNP 72

OY 83 CHNGTCEIS- EAYRGDTFEGYCKPFGNGHCHNHINECEVEPCXKNGICITDLVANY 141

Db 73 CYNDAKCLVETDQRDIFTEYICQCPVSGSGHCEETNYND- -GEYFTAVPNT 129

OY 142 SCECPGEF- MGRNCOYACGPIGIEGGLISNQITASTTRALFGLQKWPYARLNKK 201

Db 109 -----LG----- CSTRGLDEGAIADSQISASVYMGFGLQRMGPGLARLYRT 154

OY 202 INAMTAANDMKRMWIOINLQKRMVGVITQAKRIGSPYIKFYIAYSDNGRTWA 261

Db 155 VYANMAYSSYS- KFWIOVNLKRMVSGVMTGASRAGAEYLTFFVAASLDGRFE 213

OY 262 MYKVGINEMVFRGNIDNNTPYANSTPPIKAQVRLYPQVCRHRTLMELGCELSG 321

Db 214 ODESGTGKEFMGNQDNLSKINMNPTEAOYIRLYPVSCRHCTLRELLGCELSG 273

OY 322 EPLGKSGHIDVOITASSIFRTLMDFEWRKARLDKQKVNANTSGHDSQML 381

Db 274 EPLGKSGHIDVOITASSIFRTLMDFEWRKARLDKQKVNANTSGHDSQML 333

OY 382 DDLVTKVYITITQAKDGHVQFVSGKLAISNDEHWTYODEKORAKYFQGNED 441

Db 334 DLGTQKRYTITITQAKDGHVQFVSGKLAISNDEHWTYODEKORAKYFQGNED 391

OY 442 THKRVNIDPPYARIRIRILPMSVWGRTITLASELLGC 477

Db 392 SHKKNIFEKPMARYAVLPVSMNRITLRELLGC 427

RESULT 4

SW4211

A: protein precursor - bovine

C: Species: Bos taurus (cattle)

C: Date: 12-Dec-1997 #text_change 20-Jun-2000

A:Accession: SW4211; MUID:91620006; PIDN:BAH12210.1; PID:91620007

R:Stubbs, J.D.; Lekutis, C.; Singer, K.L.; Bul, A.; Yuzuki, D.; Srinivasan, U.; Parry, G

A:Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the ext

A:Reference number: A36479; MUID:91046008

A:Accession: A36479

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-427 <RVA>

A:Cross-references: EMBL:A91895; NID:91632778; PIDN:CAA62997.1; PID:91632779

A>Title: dlk pc2 and Pref-1 mRNAs encode similar proteins belonging to the EGF-like superfamily
A:Reference number: S53716; MUID:95226449

A:Molecule type: mRNA
A:Residues: 1-385 <LEP>
A:Cross-references: EMBL:U15980; NID:g562107; PIDD:AAB60495.1; PID:g562108
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: transmembrane protein
F:54-85/Domain: EGF homology <EGFP>
F:92-124/Domain: EGF homology <EGFP>
F:131-167/Domain: EGF homology <EGFP>
F:214-246/Domain: EGF homology <EGX1>
F:303-332/Domain: transmembrane #status predicted <TM1>

Query Match 9.18; Score 244.5; DB 2; Length 385;
Best Local Similarity 31.6%; Pred. No. 1.le-10;
Matches 50; Conservative 18; Mismatches 65; Indels 25; Gaps 6;

Oy 26 CDNPECGGCTGLPLAVGSFSCPCDPTDCSSVVEVASDEEPTSGACFPN--PC 83
 | | | | | : | | | | | : | | | | | : | | | | |
Db 92 CTSTPCANNNGTGV-DLEKQYECSCTPGFSFGKDOH-----RAGCVINSPC 138
 | | | | | : | | | | | : | | | | | : | | | | |

Oy 84 HNGGTETSEARRGDTFIGYVCCKPRGFNGHCQ--HNINECEVERCKNGICTLVANY 141
 | | | | | : | | | | | : | | | | | : | | | | |
Db 139 QHGCAVCDDEGASHA-----SCLPPGFSGNCETVAATNSCTPNCEMDGCTDIGDF 194
 | | | | | : | | | | | : | | | | | : | | | | |

Oy 142 SCECPGEFGRNCRCK-----CSGPLIDEGGIISNOIR 175
 | | | | | : | | | | | : | | | | | : | | | | |
Db 195 KRCRPAGFYDKTCRPVSNCSASGPCONGTCLQHNVQS 232
 | | | | | : | | | | | : | | | | | : | | | | |

RESULT 13

A54785
Preadipocyte factor 1 precursor, long form - mouse
M:Alternate names: delta-like dlk homeotic protein, pref-1
C:Species: Mus musculus (house mouse)
C:Date: 13-Apr-1995 #sequence, revision 12-Apr-1995 #ext_c-change 11-Jan-2000
A:Accession: A54785; D43484; A40746; S21385
R:Smns,C.N.,Green,D.J.,Sul,H.S.
Biochemistry 33, 9257-9265, 1994
A>Title: Structural characterization and alternate splicing of the gene encoding the preadipocyte factor 1
A:Reference number: A54785; MUID:94325292
A:Accession: A54785

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA
A:Residues: 1-385 <SMAP>
A:Cross-references: GB:S71340
R:Laborde,A.J.; Sauvalle,E.A.; Hoffman,T.; Notario,V.
J. Biol. Chem. 268, 3817-3820, 1993
A>Title: dlk, a putative mammalian homeotic gene differentially expressed in small cell lung carcinoma
A:Reference number: A5484; MUID:93179372
A:Accession: A5484

A:Molecule type: mRNA
A:Residues: 1-78, 'G', '80-343, 'TF', 346-385 <LAB>
A:Cross-references: EMBL:Z12171; NID:g50716; PIDD:CA81862.1; PID:g50717
A>Note: sequence extracted from NCBI backbone (NCBI:125734)
A>Note: species designations for this sequence report and for B4584 originally were taken from Cell
R:Smns,C.M.; Sul,H.S.
Cell 73, 725-734, 1993
A>Title: Pref-1, a protein containing EGF-like repeats, inhibits adipocyte differentiation
A:Reference number: A40746; MUID:93272313
A:Accession: A40746

A:Molecule type: mRNA
A:Residues: 1-78, 'G', '80-249, 'P', 251-319, 'CMAMPSSFSSTAKGCCCATTTCCAPARRTSGCCSITAAWRKSISST
A:Cross-references: GB:I11221; NID:g309093; PIDD:AAA37175.1; PID:g309093
A:Experimental source: 3T3-L1 preadipocytes
A>Note: sequence extracted from NCBI backbone (NCBIN:132685, NCBI:P132713)
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; tandem repeat; transmembrane protein
F:1-385/Product: preadipocyte factor 1 precursor splice form A status predicted <FOA>
F:1-250,262-385/Product: preadipocyte factor 1 precursor splice form B status predicted
F:1-250,262-385/Product: preadipocyte factor 1 precursor splice form C status predicted

Query Match	Best Local Similarity	Score	DB 2:	Length	365:
Matches 50: Conservative 18: Mismatches 65: Indels 25: Gaps 6:					
QY 26 CDNPENCGICLPLGLAVGSFSCPCDFGTPDNCSSVVEASDEEPTSAQCTPN--PC 83					
Db 92 CTTPPANGNGICV-DLEKGYECSTCPGSGSDQH-----KAGCYINSSPC 138					
QY 84 HNGCTCISEAVRGDTPIGYVCKCPRGFNGIHCQ-HNINECEVFCNKNGICTDLVANY 141					
Db 139 OHGAGVADDEQASHA-----SCLCPGFSGNRCETVAATNSCTPNPCENDGCTDIDGDP 194					
QY 142 SCECPGEFGNRNCOYK---CSGPLGIEGGLIISNQDIT 175					
Db 195 RCRCPAGFYDKTCSRPNVSNCSAGSPCQNGTCLQHTQVS 232					
RESULT 14					
B49175					
Notch A protein - mouse (fragment)					
N:Alternate names: Notch homolog					
C:Dates: 21-Jan-1994 #sequence, revision 05-Jan-1996 #ext_change 20-Sep-1999					
C:Species: Mus musculus (house mouse)					
C:Accession: B49175: PH1569: S32109					
R:Lardelli, M., Lendahl, U					
Exp. Cell. Res. 204, 364-372, 1993					
A:Title: Notch A and Notch B/two mouse Notch homologues coexpressed in a wide variety					
A:Reference number: A49175: K0120:93178563					
A:Status: preliminary: nucleic acid sequence not shown					
A:Accession: B49175					
A:Molecule type: mRNA					
A:Residues: 1-387 <L&R>					
A:Cross-references: EMBL:X68278; NID:g287987; PIDN:CAA48339.1; PID:g287988					
A:Experimental source: embryo					
A:Note: sequence extracted from NCBI backbone (NCBI:P126159)					
C:Comment: This protein has many EGF repeats and 11-12/Notch repeats.					
C:Superfamily: unsigned ankryin repeat proteins controlling the decision bea					
F:27-58/Domain: EGF homology <EGF>					
F:73-104/Domain: EGF homology <EGF2>					
F:151-185/Domain: EGF homology <EGF1>					
Query Match	8.9%: Score 239: DB 2: Length 387:				
Best Local Similarity 30.3%: Pred. No. 1.1e-10:					
Matches 57: Conservative 13: Mismatches 76: Indels 42: Gaps 6:					
QY 26 CDNPENCGICLPLGLAVGSFSCPCDFGTPDNCSSVVEASDEEPTSAQCTPNCHN 85					
Db 27 CLSQPQNGGCTID--LTNSYKSCPRGNGVHCINVDCHPDLPAASRSP---KCFN 80					
QY 86 GGTGCEISEAVRGDTPIGYVCKCPRGFNGIHCQHNINCEVECPKNGSI--CTDLYAVYS 143					
Db 81 NGRCV-----DVGSGYTCTCPRGVBERREGGVNCLSNPDGRCQNGVQVYNDFFH 133					
QY 144 ECGGEPMGRNQ-----YCSPLGIEGGLIISNQDITAS 177					
Db 134 ECGAGTGRCEISYINGCRCKPCKNGVCAYASNTANGFICRCAPGFEATCENDANTGCS 193					
QY 178 STRALFG 185					
Db 194 SL-RCLNG 200					
RESULT 15					

T46247
hypothetical protein DKFP761G02121.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46247
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223029
A:Accession: T46247
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-252 <AAA>
A:Cross-references: EMBL:AL137311
A:Experimental source: adult amygdala; clone DKFP761G02121
C:Genetics:
A>Note: DKFP761G02121.1

Query Match 7.6%; Score 204.5; DB 2; Length 252;
Best local similarity 30.6%; Pred. No. 6e-08;
Matches 45; Conservative 14; Mismatches 51; Indels 37; Gaps 6;
QY 25 ICDPM-----PCENGGICLPGLAVGSPSCPCPDGFTDPNCSSVEVASD 68
Db 7 LCDPGRHGYCEBEYNECLSPCLNATCRD--LVNGTECVCLAEYKGTIC----ELYKD 60
QY 69 EEEPTISAGPCTPNPCDHNGSGTCEISEAYRGDTFTGYVCKCPRGENGHCOHNINECEVEPC 128
Db 61 -----PCANVSCNLNGATCD--SDGLNG-----TCICAPGFTGECDDIDINECDSPC 105
QY 129 KNGGICTDLVANYSCCPCGEGMGRCQ 155
Db 106 HHGGSCLDQPMNGYNCHCPHGVGANCE 132

Search completed: May 23, 2001, 08:38:48
Job time: 109 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:38:09 ; Search time 11.45 Seconds
(without alignments)
1439.030 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513
2675
Sequence: 1 MKRSVAVMLVGLSLGVPOF.....MSWYGRITLASELCTREE 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 segs, 34255486 residues
Total number of hits satisfying chosen parameters: 71834

Minimum DB seq length: 0
Maximum DB seq length: 481

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1173	43.9	409	1	MEGM_PIG
2	1156	43.2	463	1	MEGM_MOUSE
3	1154	43.1	427	1	MEGM_RAT
4	1143	42.7	427	1	MEGM_BOVIN
5	1020.5	38.1	387	1	MEGM_HUMAN
6	295.5	11.0	280	1	XLRI_FIGRU
7	280	10.5	224	1	XLRI_MOUSE
8	268	10.0	473	1	FP2_MYTGA
9	267	10.0	224	1	XLRI_HUMAN
10	245	9.2	383	1	DLK_MOUSE
11	244.5	9.1	385	1	WFI1_MOUSE
12	185.5	6.9	379	1	WFI1_HUMAN
13	182	5.6	448	1	PR15_MOUSE
14	150	5.6	448	1	PR15_MOUSE
15	149	5.6	448	1	PR15_MOUSE
16	148.5	5.6	407	1	PR2_BOVIN
17	148	5.5	475	1	PR2_BOVIN
18	146	5.5	448	1	PR2_BOVIN
19	140.5	5.3	257	1	PR2_BOVIN
20	135	5.0	444	1	PR2_BOVIN
21	134.5	5.0	444	1	PR2_BOVIN
22	134.5	5.0	444	1	PR2_BOVIN
23	131	4.9	470	1	PR2_BOVIN
24	131	4.9	470	1	PR2_BOVIN
25	126.5	4.7	396	1	PR2_BOVIN
26	125	4.7	396	1	PR2_BOVIN
27	125	4.7	396	1	PR2_BOVIN
28	123.5	4.6	443	1	PR2_BOVIN
29	123	4.6	443	1	PR2_BOVIN
30	123	4.6	443	1	PR2_BOVIN
31	123	4.6	443	1	PR2_BOVIN
32	121.5	4.5	253	1	PR2_BOVIN
33	121.5	4.5	253	1	PR2_BOVIN

ALIGNMENTS

RESULT	ID	MEGM_PIG	STANDARD	PRT	409 AA.
AC	MEGM_PIG	01-NOV-1997 (Rel. 35, Created)			
DT	15-JUL-1998 (Rel. 39, Last annotation update)				
DT	30-MAR-2000 (Rel. 39, Last annotation update)				
DE	LACTADHERIN (MILK FAT GLOBULE-EGF FACTOR 8) (MEG-8) (MEGM) (SPERM SURFACE PROTEIN SP47) (PP47).				
GN	MEG8.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.				
OX	NCBI_TaxID=9823;				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Testis;				
RA	Ensembl M.A.:				
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.				
CC	- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-BINDING PROTEIN.				
CC	- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.				
CC	- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND SPERMATOZOA.				
CC	- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.				
CC	- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
DR	EMBL; Y11683; CAAT7379.1; -				
DR	HSSP; P00740; IIXA.				
DR	InterPro; IPR000421; -				
DR	InterPro; IPR000561; -				
DR	Pfam; PF00008; EGF_2.				
DR	Pfam; PF00754; F5-F8 type C; 2.				
DR	PROSITE; PS01022; EGF_2.				
DR	PROSITE; PS01186; EGF_2.				
DR	PROSITE; PS01285; F5B6C_1; 2.				
DR	PROSITE; PS01286; F5B6C_2; 2.				
KW	Glycoprotein; Repeat; EGF-like domain.				
FT	DOMAIN	2	41		
FT	DOMAIN	44	88		
FT	DOMAIN	91	247		
FT	DOMAIN	252	409		
FT	SITE	67	69		
FT	DISULFID	6	17		
FT	DISULFID	11	29		
FT	DISULFID	31	40		
FT	DISULFID	91	247		
FT	DISULFID	234	238		
FT	DISULFID	234	238		

FT DISULFID 252 409 BY SIMILARITY.
 FT CARBOHYD 41 41 N-LINKED (GLCNC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNC. . .) (POTENTIAL).
 SQ SEQUENCE 409 AA: 45725 MW: 80C07AF80029927A CRC64;

Query Match 43.9%: Score 1173; DB 1; Length 409;
 Best Local Similarity 48.0%: Pred. No. 8 2e-82;
 Matches 223; Conservative 63; Mismatches 111; Indels 68; Gaps 8;

23 GDICDPNCPENGICL-----PGLAVSFCSCPDGCTDPCNGSVYEVASDEEPTSA 75
 DB 3 GDCDSSLCLNGICLTLDQDPK-----FRLCESTFTGLICNE-----TER 45
 QY 76 GPCFPNCPHNGGTCF-ISEAVRGCTFGYVCKCPGFCNGICQHNNINCEVEYPCCKMGIC 134
 DB 46 GPCFPNCPHNDACEYIDAHRGVFTFYICKCPHGTGIR----- 86
 QY 135 TDVANYSCDCEPGEFMRNCOYKCSGPLGIEGSIISNDOITASSTHRLGLOKWPYYA 194
 DB 87 -----CETLONAPLGMETGALADFGIISMSHMGFOLRMWAPELA 127
 QY 195 RLKKGLINAWTAENDRMKRMIOINDRMKRVGYITOGAKRIGSEPEYIKFIAYSD 254
 DB 128 RLHRAGIVAWTASNYDR-NPMIOVNLIRRMRYGVVTOGASRAGSAEYMKTFVAYSTD 186
 QY 255 GKTWAMKVKGTNE--DWVFRGNDNTNTPYANSFTPIKAOYRVLRYPOVCRHCTIRFEL 312
 DB 187 GRKFOF--IOGAEESGDKIFMGINDNSGLVNLFEVPLEVOYVRLVPLICRGTIRFEL 244
 QY 313 LGCELSGSEPLGMSGHIDYOTASTIRPTNMMFWERKARADKQKNAWTSQH 372
 DB 245 LGCELSGCAEPGLKMDNTIPMKOITASFTRTGSLASFWPFAKIDNCKRNMTAS 304
 QY 373 NDSQMLVDLVNPKYKGIITOGAKDPGVCYGVSCYKAVSNGCEHWYVYODEKQRK 432
 DB 305 NSASEHQLDLDGQRKRVGITOGAKRFGIYVAVKAVASDDVSTFRDQALSK 364
 QY 433 VFGGNDNTNHRKKNVDPPIYAHIRILPMSWGRITLASELGC 477
 DB 365 IFPGNDNNSHKKNMFETPLIRFRLPVAHNRITLVELLGC 409

RESULT 2
 MFGM_MOUSE STANDARD: PRT: 463 AA.
 ID MFGM_MOUSE
 AC P21956; P97800;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-EB) (MFGM)
 DE (SPERM SURFACE PROTEIN SP47) (MP47).
 GN MFG8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-35.
 RC TISSUE=Mammary gland.
 RX MEDLINE=9104608; PubMed=2122462;
 RA Studts J.D., Lekuts C., Singer K.L., Bul A., Yuzuki D.,
 RA Stinivasan U., Parry G.;
 RT "cDNA cloning of a mouse mammary epithelial cell surface protein
 RT reveals the existence of epidermal growth factor-like domains linked
 RT to factor VIII-like sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
 RN [2]
 RP SEQUENCE OF 23-463 FROM N.A.
 RC TISSUE=Testis;
 RA Ensslin M.A.;
 RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-

CC BINDING PROTEIN.
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND
 CC SPERMATOZOAN.
 CC -1- DEVELOPMENTAL STAGE: mRNA EXPRESSION IS DETECTABLE IN MAMMARY
 CC TISSUE FROM NONPREGNANT ANIMALS & MAXIMAL IN THE LACTATING GLAND.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 FS/8 TYPE C DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M38337; AAA39534.1; -;
 CC EMBL: Y11684; CAA72380.1; -;
 CC PIR: A36479; A36479.
 CC HSSP: P00740; 11XA.
 CC MCD: MGI:102768; Mf9e8.
 CC Interpro: IPR000421; -;
 CC Interpro: IPR000561; -;
 CC Interpro: IPR001438; -;
 CC Pfam: PF00008; EGF_2.
 CC Pfam: PF00754; F5_F8_Type_C_2.
 CC PRINTS: PR00010; EGFRIOD.
 CC PROSITE: PS00022; EGF_1; 2.
 CC PROSITE: PS01186; EGF_2; 2.
 CC PROSITE: PS01285; FR8C_1; 2.
 CC PROSITE: PS01286; FR8C_2; 2.
 CC KAM: signal; glycoprotein; Repeat; EGF-like domain; MILK.
 CC SIGNAL 1 2
 CC CHAIN 2 2
 CC DOMAIN 24 61
 CC DOMAIN 64 108
 CC DOMAIN 148 303
 CC DOMAIN 308 463
 CC SITE 87 89
 CC SITE 87 89
 CC DISULFID 33 49
 CC DISULFID 51 60
 CC DISULFID 68 79
 CC DISULFID 73 96
 CC DISULFID 98 107
 CC DISULFID 148 303
 CC DISULFID 290 463
 CC DISULFID 308 463
 CC CARBOHYD 61 61
 CC CARBOHYD 266 266
 CC CARBOHYD 316 316
 CC CARBOHYD 426 426
 CC CARBOHYD 436 436
 CC CONFLICT 30 30
 CC CONFLICT 35 35
 CC CONFLICT 110 147
 CC CONFLICT 168 168
 CC CONFLICT 196 196
 CC CONFLICT 309 309
 CC CONFLICT 395 395
 CC SEQUENCE 463 AA: 51465 MW: D78B6C6EFBA724D CRC64;

Query Match 43.2%: Score 1156; DB 1; Length 463;
 Best Local Similarity 48.5%: Pred. No. 1.8e-80;
 Matches 222; Conservative 72; Mismatches 142; Indels 22; Gaps 8;

23 GDICDPNCPENGICLPLAVSFCSCPDGCTDPCNGSVYEVASDEEPTSA 82
 DB 25 GDCDSSLCLNGICLTLDQNDIYLCPEPGFTGLVNC-----TERGSPSPRP 72
 QY 83 CHNGGICEIS-EATYNGDTFTIYVCKCPGFCNGICQHNNINCEVEYPCCKNGICTDVANY 141

```

DR 73 CYNDAKCLVLTDPGRGDIETREYICOPVGSIGCHETETNYULD--GEYMTAVANT 129
QY 142 SECEGGER--MGRMCQKCSGPIEGGIISNOOITSSPHRALPFLQKMYVYARLANK 199
DB 130 AVPTAPRPDLNNLASRCSTOLGMEGALDSQISLSYVYMFQMLQWELARLRYT 169
QY 200 GLINMTAENDRKRMIOINLQRMKRVGYITOGARISPRYKFIYLAISDQKTM 259
DB 190 GIVNMMHNSYDS-KPMIOVNLKRRSGVMTQASRARARLKTETFAVSLDGRFTE 248
QY 260 MYKVGITNEDVFRGNIDNNTPYANSFTPIKAQVYLYQVCRHCTLRMELLCESG 319
DB 249 FTIDE-SGDKKEFLGNDNNSLKVNPFTLEAOYIRLYVPSCHRCCTLRPELLGCELHG 307
QY 320 CSEPLGKMSGHIODYQITASSIFRTLNMDFTWEPKARLDKQKVNANTSGHDSQWL 379
DB 308 CLEPLGKRNNTIPDSOMASSSYKTMNLRAGWYPHLGRDNOGKINAMTAOSNKAEML 367
QY 380 QVLLVPTKVTGIIITOGAKDFGHVQFVGSYKLAYSNDGEHMTVYODEKORCKDPQGNFD 439
DB 368 QVLDGTQROVYGIITOGARDFGHIOVESYKVAHSDGQVOMTVY--EEGSSKVPQGNLD 425
QY 440 NDTHRKNVIDPPIYARHRIPLPMSWYGRITLASELLGC 477
DB 426 NNSHKNIPEKPFMARVYVLPVSMHNRITLRLELLGC 463

RESULT 3
MSGM RAT STANDARD: PRT: 427 AA.
AC P70490:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (O-
DE ACETYL GD3 GANGLIOSIDE SYNTHASE) (AGS) (MFGM).
GN MFGS OR AGS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI-TaxId=10116;
RN 11)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96374422; Pubmed=8780713;
RA Ogura K., Nara K., Watanabe Y., Kohno K., Tai T., Sanai Y.;
RT "Cloning and expression of cDNA for O-acetylation of GD3
RT ganglioside."
RL Biochem. Biophys. Res. Commun. 225:932-938(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. SEEMS TO
CC PARTICIPATE IN THE O-ACETYLTATION OF GD3 GANGLIOSIDE SIALIC ACID.
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: SPLEEN, LUNG, HEART, BRAIN AND MUSCLE.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
CC EMBL: DB4068; BA12210.1; -
CC HSSP: P00740; ILYA.
CC InterPro: IPR000421; -
CC InterPro: IPR000561; -
CC InterPro: IPR00138; -
CC Pfam: PF00008; EGF_2.
CC Pfam: PF00734; F5_Type_C_2.
CC PRINTS: PR00010; EGFBL00D.

```

```

DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
KW Signal; glycoprotein; Repeat; EGF-like domain; Milk.
FT SIGNAL 1 22
FT CHAIN 23 427
FT DOMAIN 24 61 EGF-LIKE 1.
FT DOMAIN 64 108 EGF-LIKE 2.
FT DOMAIN 111 267 F5/8 TYPE C 1.
FT DOMAIN 272 427 F5/8 TYPE C 2.
FT DISULFD 28 39 BY SIMILARITY.
FT DISULFD 31 49 BY SIMILARITY.
FT DISULFD 51 60 BY SIMILARITY.
FT DISULFD 68 79 BY SIMILARITY.
FT DISULFD 98 107 BY SIMILARITY.
FT DISULFD 111 267 BY SIMILARITY.
FT DISULFD 254 258 BY SIMILARITY.
FT DISULFD 272 427 BY SIMILARITY.
FT SITE 87 89 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 427 AA; 47413 MW; EA8C8631F3EE6047 CRC64;

```

Query Match 43.1%; Score 1154; DB 1; Length 427;
 Best local similarity 48.5%; Pred. No. 2,4e-80;
 Matches 221; Conservative 62; Mismatches 119; Indels 54; Gaps 7;

```

QY 23 GDIQPNPCNMGKICGLGLAVSFGSCPDGFTDPCNSVVEVASDEEPTSGAGCTPNP 82
DB 25 GDFCSLCLNGGTCIMG-QDNDIYCLCPREFTGLVONE-----TEKGPCSPNP 72
QY 83 CANGTCEISE-AARGTEFLGYCKCPRENGJHCCHININECEVERCKNGGICDVLVANY 141
DB 73 CRDPAKLYTEDTORSHDITREYICOPVGSIGCH----- 108
QY 142 SECEGGERMGNQKCSGPIEGGIISNOOITSSPHRALPFLQKMYVYARLANKGL 201
DB 109 -----LG-----CYTLGEGSALDSQISLSYVYMFQMLQWELARLRYGI 154
QY 202 INAMTAENDRKRMIOINLQRMKRVGYITOGAKRIGSEPTIKFIKIASDQKTMAY 261
DB 155 VAMTAASSYDS-KPMIOVNLKRRSGVMTQASRARARLKTETFAVSLDGRFPEFI 213
QY 262 KYKGTNEDVFRGNIDNNTPYANSFTPIKAQVYLYQVCRHCTLRMELLCESGS 321
DB 214 ODESGTGKREMGNDNNSLKVNPFTLEAOYIRLYVPSCHRCCTLRPELLGCELHG 273
QY 322 EPLGKMSGHIODYQITASSIFRTLNMDFTWEPKARLDKQKVNANTSGHDSQWL 381
DB 274 EPLGKRNNTIPDSOITASSSYKTMNLRAGWYPHLGRDNOGKINAMTAOSNKAEML 333
QY 382 DLVLVPTKVTGIIITOGAKDFGHVQFVGSYKLAYSNDGEHMTVYODEKORCKDPQGNFD 441
DB 334 DLTGQKVTGIIITOGARDFGHIOVASVYKVAHSDGQVOMTVY--EEGSSKVPQGNLD 391
QY 442 THRKKNVIDPPIYARHRIPLPMSWYGRITLASELLGC 477
DB 392 SHKNIPKPFMARVYVLPVSMHNRITLRLELLGC 427

RESULT 4
MSGM BOVIN STANDARD: PRT: 427 AA.
AC 095114; 027959; P79344;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8)

```



```

OS Eukaryotes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCB1_TaxID=31033;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=9263230; PubMed=10330123;
RA Brunner B., Todt T., Lenzner S., Stout K., Schultz U., Ropers H.-H.,
RA Kalscheuer V.M.;
RT "Genomic structure and comparative analysis of nine Fugu genes:
RT conservation of synteny with human chromosome Xp22.2-p22.1.";
RL Genome Res. 9:437-448(1999).
CC -1- FUNCTION: MAY BE ACTIVE IN CELL ADHESION PROCESSES DURING RETINAL
CC DEVELOPMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 FS/8 TYPE C DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF146687; AAD8797.1; -
DR InterPro: IPR00421; -
DR Pfam: PF00754; F5_F8_Type_C_1.
DR PROSITE: PS01285; F558C_1; 1
DR PROSITE: PS01286; F558C_2; FALSE_NEG.
KW Cell adhesion; Signal.
FT SIGNAL 1
FT CHAIN 1
FT DOMAIN 119 280
FT DISUFID 119 275
FT SEQUENCE 280 AA; 32032 MW; ADF9A3222ED0167 CRC64;

Query Match 11.0%; Score 295.5; DB 1; Length 280;
Best Local Similarity 42.2%; Pred. No. 1.0e-15;
Matches 68; Conservative 24; Mismatches 54; Indels 15; Gaps 5;

QY 333 PLKMSGHTQDYQITASSIRTLANDMT-----WEPRKARDKQGVNAWTSQHNDOSQ 377
DB 124 PLGFEGSISPDQITCS-----NQOQITAMPSSMLSKRLMTQSGFCAMLSKFDQNTQ 177
QY 378 WLQVDLVPYTKVGTITGAKDGHVQVSGYKTAISNDE-HWIVYQDEKQKDKYFQG 436
DB 178 WLQDLDLDAKVSGILVGRCDAD--EWITKISLQYRDKRLMTWITTKDQ-TGNRRVYFG 234
QY 437 NFDNDTHRKKNVDPPIYARHRIPLPMSWGRITLASLLECC 477
DB 235 NSDRSSVQNLRLPPIYARHRIPLPMSWGRITLASLLECC 275

RESULT 7
XRL_MOUSE STANDARD; PRT: 224 AA.
ID XRL_MOUSE
AC Q921L4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE X-LINKED JUVENILE RETINOSCHISIS PROTEIN PRECURSOR.
GN RSI OR RSIH OR XLRSI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RF STRAIN=129/SVJ;

```

```

RX MEDLINE=99160474; PubMed=10051329;
RA Gehrig A.E., Warneke-Mittstock R., Sauer C.G., Weber B.H.F.;
RT "Isolation and characterization of the murine X-linked juvenile
RT retinoschisis (Rsh) gene.";
RL Mamm. Genome 10:303-307(1999).
RN (12)
RP SEQUENCE FROM N.A.
RX STRAIN=C57/BL6; TISSUE=Eye;
RA Kalscheuer V.M.;
RT "Genomic structure and comparative analysis of seven contiguous genes
RT disclose a large region with conserved gene order in human Xp22.2-
RT p22.1.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN (13)
RP SEQUENCE FROM N.A.
RX TISSUE=Retina;
RX MEDLINE=99148018; PubMed=10023077;
RA Reid S.N., Akhmedov N.B., Pirtlev N.I., Kozak C.A., Danciger M.,
RA Farber D.B.;
RT "The mouse X-linked juvenile retinoschisis cDNA: expression in
RT photoreceptors.";
RL Gene 227:257-266(1999).
CC -1- FUNCTION: MAY BE ACTIVE IN CELL ADHESION PROCESSES DURING RETINAL
CC DEVELOPMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: RETINA-SPECIFIC.
CC -1- SIMILARITY: CONTAINS 1 FS/8 TYPE C DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF084561; AAD21808.1; -
DR EMBL: AF084561; AAD21809.1; -
DR EMBL: AF084562; AAD21809.1; JOINED.
DR EMBL: AF084563; AAD21809.1; JOINED.
DR EMBL: AF084564; AAD21809.1; JOINED.
DR EMBL: AF084565; AAD21809.1; JOINED.
DR EMBL: AF084566; AAD21809.1; JOINED.
DR EMBL: AF011381; CAA09601.1; -
DR MSG: MG1:1336189; Rsh.
DR InterPro: IPR00421; -
DR Pfam: PF00754; F5_F8_Type_C_1.
DR PROSITE: PS01285; F558C_1; 1.
DR PROSITE: PS01286; F558C_2; FALSE_NEG.
KW Cell adhesion; Signal.
FT SIGNAL 1
FT CHAIN 1
FT DOMAIN 63 224
FT DISUFID 63 219
FT SEQUENCE 224 AA; 25575 MW; 4536203C0D0E90E4 CRC64;

Query Match 10.5%; Score 280; DB 1; Length 224;
Best Local Similarity 33.8%; Pred. No. 1.0e-14;
Matches 67; Conservative 33; Mismatches 74; Indels 24; Gaps 7;

QY 298 YPOVCRNRHCTLMELL--CGLSGC-----SEPLKMSGHTQDYQITASSIRTLAND 348
DB 34 YORACDCDQGVANALMSAGATSLDICEPCYHRLPFESGEVTPQDITCS-----NPE 87
QY 349 MF-----TWEPKARDKQGVNAWTSQHNDOSQWLQVDLVPYTKVGTITGAKDGHV 403
DB 88 QYGVMSWTKARLRKRLMSQSGFCAMLSKXQDSQWLQIDKEIKVYSGILVGRCDID-- 145
QY 404 QYGVSKTAISNDE-HWIVYQDEKQKDKYFQGNFNDTHRKKNVDPPIYARHRIPLP 462
DB 146 EHWTKSYQVYRDLRLMWITTKDQ-TGNRRVYFGNSDRSSVQNLRLPPIYARHRIPL 204

```


RESULT	8
FP2_MITGA	STANDARD; PRG: 473 AA.
ID	FP2_MITGA
AC	Q23464;
DY	30-MAY-2000 (Rel. 39, Created)
DY	30-MAY-2000 (Rel. 39, Last sequence update)
DY	01-OCT-2000 (Rel. 40, Last annotation update)
DE	ADHESIVE PLAQUE MATRIX PROTEIN 2 PRECURSOR (FOOT PROTEIN 2) (MGFP2)
DE	(MGFP-2).
GN	FP2.
OS	Mytilus galloprovincialis (Mediterranean mussel).
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC	Mytiloidea; Mytilidae; Mytilus.
OX	NCBI_TaxID:29158;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Foot;
RX	MEDLINE:95204464; PubMed:7996812;
RT	Inoue K., Tateuchi Y., Miki D., Odo S.;
RT	"Mussel associated plaque protein gene is a novel member of epidermal
RT	growth factor-like family".
RL	J Biol. Chem. 270:6698-6701(1995).
CC	-1- FUNCTION: PROVIDES ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSELS
CC	PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S
CC	ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A
CC	FIBROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.
CC	-1- SUBCELLULAR LOCATION: SECRETED.
CC	-1- TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.
CC	-1- DEVELOPMENTAL STAGE: EXPRESSION STARTS AT THE PEDIVELGER, FOOT
CC	FORMATION, STAGE.
CC	-1- PTM: SOME TYROSINES ARE HYDROXYLATED (THUS PRODUCING DOPA = 3,4-
CC	DIHYDROXYPHENYLALANINE).
CC	-1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license@ebi-sib.ch) (See http://www.isb-sib.ch/announce/
CC	or send an email to license@ebi-sib.ch).
CC	-----
DR	EMBL: DA3794; BAA07852.1; -
DR	InterPro: IPR000561; -
DR	Pfam: PF00008; EGF_11.
DR	PROSITE: PS00010; ASX_HYDROXYL_2.
DR	PROSITE: PS00022; EGF_1; 11.
DR	PROSITE: PS01186; EGF_2; 10.
DR	GlycoProtol: EGF-like domain; Repeat; Signal; Hydroxylation.
FT	SIGNAL
FT	CHAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	DOMAIN
FT	MOD_RES
FT	MOD_RES
FT	MOD_RES
FT	CARBOHYD

FT	DISULFID	49	60		BY SIMILARITY.
FT	DISULFID	54	69		BY SIMILARITY.
FT	DISULFID	71	80		BY SIMILARITY.
FT	DISULFID	86	97		BY SIMILARITY.
FT	DISULFID	91	106		BY SIMILARITY.
FT	DISULFID	108	117		BY SIMILARITY.
FT	DISULFID	122	133		BY SIMILARITY.
FT	DISULFID	127	143		BY SIMILARITY.
FT	DISULFID	145	154		BY SIMILARITY.
FT	DISULFID	159	170		BY SIMILARITY.
FT	DISULFID	164	180		BY SIMILARITY.
FT	DISULFID	182	191		BY SIMILARITY.
FT	DISULFID	196	207		BY SIMILARITY.
FT	DISULFID	201	217		BY SIMILARITY.
FT	DISULFID	219	228		BY SIMILARITY.
FT	DISULFID	233	244		BY SIMILARITY.
FT	DISULFID	238	254		BY SIMILARITY.
FT	DISULFID	256	265		BY SIMILARITY.
FT	DISULFID	270	281		BY SIMILARITY.
FT	DISULFID	275	290		BY SIMILARITY.
FT	DISULFID	292	301		BY SIMILARITY.
FT	DISULFID	306	317		BY SIMILARITY.
FT	DISULFID	311	328		BY SIMILARITY.
FT	DISULFID	330	339		BY SIMILARITY.
FT	DISULFID	346	357		BY SIMILARITY.
FT	DISULFID	351	366		BY SIMILARITY.
FT	DISULFID	368	377		BY SIMILARITY.
FT	DISULFID	387	399		BY SIMILARITY.
FT	DISULFID	393	408		BY SIMILARITY.
FT	DISULFID	410	419		BY SIMILARITY.
FT	DISULFID	429	440		BY SIMILARITY.
FT	DISULFID	434	449		BY SIMILARITY.
FT	DISULFID	451	460		BY SIMILARITY.
SQ	SEQUENCE	473 AA;	51772 MW;	BAT6BAC3BA49A0F CRC64;	
Query Match 10.0% Score 268; DB 1; Length 473;					
Best Local Similarity 37.6%, Pred. No. 3,4e-13;					
Matches 59; Conservative 17; Mismatches 53; Indels 28; Gaps 8;					
OY	20 FGR-----GDICDPCNCGNGGCIPLGLAVGSFSFGDPFPNCSSVVFVASDSEEPISAG 76				
DB	113 FGFLCKKNKCVSPFKNNKGCSF-LKRTSKTKCTGSGGYTPRC-----EVHA----- 158				
OY	77 PCGPFCNHGNGTCEISEARYNGDTIFIGIVCKCPKPGFNHCOHNINCEVEPECKNGICT- 135				
DB	159 -CKPNCCKKKKGC-----FPDKTKTKKRCVDGSGPTOE--NACKPNCSNGGTCSA 209				
OY	136 DLYANTSCCEPGEPMGNCO-YKCSGPLGIBGGIISN 171				
DB	210 DKFGDYSCBCRGFYGECERCERYCACPNCKNGIGCSS 246				
RESULT 9					
ID	XLR1_HUMAN	STANDARD;	PRT;	224 AA.	
AC	O15537;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	X-LINKED JUVENILE RETINOSCHISIS PROTEIN PRECURSOR.				
DN	R61 OR XLR1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	KRBL-TaxId=9606;				
RN	[1] SEQUENCE FROM N.A., AND VARIANTX XLR1 ARG-96 AND TRP-102.				
RP	TISSUE=Retina;				
RA	MEDLINE=97467726; PubMed=9326935;				
RA	Sauer C.G.; Gehrig A.; Warneke-Wiltstock R.; Marguardt A.; Ewing C.C.,				
RT	Gibson A.; Lorenz B.; Jurkiles B.; Weber B.H.;				
	Positional Cloning of the gene associated with x-linked juvenile				


```

FT  VARIANT 136 136 /FTID-VAR.008226.
FT  VARIANT 138 138 I -> T (IN XLRSL1).
FT  VARIANT 140 140 /FTID-VAR.008227.
FT  VARIANT 140 140 T -> A (IN XLRSL1).
FT  VARIANT 140 140 /FTID-VAR.008228.
FT  VARIANT 140 140 G -> E (IN XLRSL1).
FT  VARIANT 140 140 /FTID-VAR.008229.
FT  VARIANT 141 141 G -> R (IN XLRSL1).
FT  VARIANT 141 141 /FTID-VAR.008230.
FT  VARIANT 141 141 R -> C (IN XLRSL1).
FT  VARIANT 141 141 /FTID-VAR.008231.
FT  VARIANT 141 141 R -> G (IN XLRSL1).
FT  VARIANT 141 141 /FTID-VAR.008232.
FT  VARIANT 141 141 R -> H (IN XLRSL1).
FT  VARIANT 142 142 /FTID-VAR.008233.
FT  VARIANT 143 143 C -> W (IN XLRSL1).
FT  VARIANT 143 143 /FTID-VAR.008234.
FT  VARIANT 146 146 D -> V (IN XLRSL1).
FT  VARIANT 146 146 /FTID-VAR.008235.
FT  VARIANT 146 146 E -> D (IN XLRSL1).
FT  VARIANT 146 146 /FTID-VAR.008236.
FT  VARIANT 146 146 E -> K (IN XLRSL1).
FT  VARIANT 155 155 /FTID-VAR.008237.
FT  VARIANT 155 155 Y -> C (IN XLRSL1).
FT  VARIANT 158 158 /FTID-VAR.008238.
FT  VARIANT 158 158 D -> N.
FT  VARIANT 163 163 /FTID-VAR.008239.
FT  VARIANT 163 163 W -> C (IN XLRSL1).
FT  VARIANT 178 178 /FTID-VAR.008240.
FT  VARIANT 178 178 G -> D (IN XLRSL1).
FT  VARIANT 182 182 /FTID-VAR.008241.
FT  VARIANT 182 182 R -> C (IN XLRSL1).
FT  VARIANT 192 192 /FTID-VAR.008242.
FT  VARIANT 192 192 P -> R (IN XLRSL1).
FT  VARIANT 192 192 /FTID-VAR.008243.
FT  VARIANT 192 192 P -> S (IN XLRSL1).

Query Match 10.0% Score 267; DB 1; Length 224;
Best Local Similarity 33.28; Pred No. 1.8e-13;
Matches 66; Conservative 34; Mismatches 73; Indels 26; Gaps 8;

DB 348 DMF-----TWPRKARLDKCKVAMNTSGHNDOSQMLQVLLVPTKVGIITGAGDFGH 402
DB 87 EGVTVGWTSSMTANKARLINSQFGCAMLKSFQSSQMLQIDKEIKVISGLITGRCDDP- 145
DB 403 VOEVGSYKLAISNDGE-HWTVYDEPKORCKVFGFNPDTHRKNNVIDPPIYARHRIILP 461
DB 146 -EMWTKYSVQYRTDERLNMWIIYKDO-TGNMRYVYGNSDRSTVONLRLPPIISRIRLIP 203
DB 462 MSWYGRITLASSELLGTEE 480
DB 204 LGMHVRALIRMLLECVSK 222

RESULT 10
D.L. HUMAN STANDARD; PRT; 383 AA.
AC PR0370;
DC 01-NOV-1995 (Rel. 32, Created)
DC 01-FEB-1996 (Rel. 33, Last sequence update)
DC 01-OCT-2000 (Rel. 40, Last annotation update)
DE DELTA-LIKE PROTEIN PRECURSOR (DLK) [CONTAINS: FETAL ANTIGEN 1 (FAL)].
GN DLK1 OR DLK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE-Adrenal gland;
RX MEDLINE=93179372; PubMed=8095043;
RA Laborda J., Sausville E.A., Hoffman T., Notario V.;
RT "dlk", a putative mammalian homeotic gene differentially expressed in
RL small cell lung carcinoma and neuroendocrine tumor cell line.";
J. Biol. Chem. 268:3817-3820(1993).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE-placenta, and Adrenal gland;
RX MEDLINE=95226449; PubMed=7711066;
RA Lee Y.L., Helman L., Hoffman T., Laborda J.;
RT "dlk", pc2 and Pref-1 mRNAs encode similar proteins belonging to the
RL EGF-like superfamily. Identification of polymorphic variants of this
RN RNA.";
RL Biochim. Biophys. Acta 1261:223-232(1995).
[3]
RN SEQUENCE OF 24-383.
RC TISSUE-embryonic fluid;
RX MEDLINE=95010145; PubMed=7925474;
RA Jensen C.H., Krogh T.N., Hoejrup P., Clausen P.P., Skjodt K.,
RA Larsson L., Englund J., Teisner B.,
RT "Protein structure of fetal antigen 1 (FAL). A novel circulating
RL human epidermal growth factor-like protein expressed in
RT neuroendocrine tumors and its relation to the gene products of dlk
RL and pc2.";
RL Eur. J. Biochem. 225:83-92(1994).
[4]
RN SEQUENCE OF 24-60.
RC TISSUE-embryonic fluid;
RX MEDLINE=93273893; PubMed=8501199;
RA Jensen C.H., Teisner B., Hoejrup P., Rasmussen H.B., Madsen O.D.,
RA Nielsen B., Skjodt K.;
RT "Studies on the isolation, structural analysis and tissue localization
RL of fetal antigen 1 and its relation to a human adrenal-specific cDNA,
RL pc2.";
RL Hum. Reprod. 8:635-641(1993).
CC -1- FUNCTION: MAY HAVE A ROLE IN NEUROENDOCRINE DIFFERENTIATION.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: A LONG FORM (SHOWN
CC HERE) AND A SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: FAL IS FOUND WITHIN THE STROMAL CELLS IN CLOSE
CC CONTACT TO THE VASCULAR STRUCTURE OF PLACENTAL VILLI, YOLK SAC,
CC FETAL LIVER, ADRENAL CORTEX AND PANCREAS AND IN THE BETA CELLS OF
CC THE ISLETS OF NEUROENDOCRINE LUNG TUMOR TISSUE.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: Z12172; CAI78163.1;
DR EMBL: U15979; AAA75364.1;
DR EMBL: U15981; AAA75365.1;
DR HSSP: P00750; 1TRG.
DR MTM: 176290;
DR InterPro: IPR000152;
DR InterPro: IPR000561;
DR Pfam: PR00008; EGF_6;
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00022; EGF_2; 5.
DR PROSITE: PS01186; EGF_1; 5.
DR Antigen: Transmembrane; EGF-like domain; Repeat; Glycoprotein; Signal;
KW Alternative splicing;
FT SIGNAL 1 23
FT CHAIN 24 383 DELTA-LIKE PROTEIN.
FT CHAIN 24 303 FETAL ANTIGEN 1.

```

```

FT DOMAIN 24 303 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 304 327 POTENTIAL.
FT DOMAIN 328 383 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 24 55 EGF-LIKE 1.
FT DOMAIN 53 86 EGF-LIKE 2.
FT DOMAIN 88 125 EGF-LIKE 3.
FT DOMAIN 127 168 EGF-LIKE 4.
FT DOMAIN 170 206 EGF-LIKE 5.
FT DOMAIN 208 245 EGF-LIKE 6.
FT DISULFID 26 37 BY SIMILARITY.
FT DISULFID 30 43 BY SIMILARITY.
FT DISULFID 45 54 BY SIMILARITY.
FT DISULFID 57 68 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 76 85 BY SIMILARITY.
FT DISULFID 92 103 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 156 BY SIMILARITY.
FT DISULFID 158 167 BY SIMILARITY.
FT DISULFID 174 185 BY SIMILARITY.
FT DISULFID 179 194 BY SIMILARITY.
FT DISULFID 196 205 BY SIMILARITY.
FT DISULFID 212 223 BY SIMILARITY.
FT DISULFID 217 233 BY SIMILARITY.
FT DISULFID 235 244 BY SIMILARITY.
FT CARBOHYD 94 94 O-LINKED.
FT CARBOHYD 100 100 O-LINKED (GLCNAC. . .).
FT CARBOHYD 143 143 O-LINKED.
FT CARBOHYD 163 163 O-LINKED (PARTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (PARTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (PARTIAL).
FT CARBOHYD 214 214 O-LINKED.
FT CARBOHYD 222 222 O-LINKED (PARTIAL).
FT CARBOHYD 251 251 O-LINKED (PARTIAL).
FT CARBOHYD 260 260 O-LINKED (PARTIAL).
FT CARBOHYD 229 301 MISSING (IN SHORT ISOPORN).
FT VARSPLIC 347 347 MISSING (IN CLONE HDLKAAG).
FT VARIANT OP--VAR.002274.
FT CONFLICT OP--HY (IN REF. 1).
FT CONFLICT G--D (IN REF. 3).
FT SEQUENCE 383 AA; 41143 MM; 2C53CE353B4415B CRC64;

```

Query Match 9.28; Score 245; DB 1; Length 383;
 Best Local Similarity 27.68; Pred. No. 1.5e-11;
 Matches 53; Conservative 24; Mismatches 59; Indels 56; Gaps 7;

```

OY 23 GDCDPN-----PCENGICLPGLAVSFCSCCPDGFDPNCSSVVEVASDEEPTSA 75
DB 82 GELCDRDVRCASAPCANNGTCV-SLDGGLYECSCAGYSGKDCOK-----KD 128
OY 76 GRCPTN--PCNNGGTCISEFAYNGDTFGLYCKPRFNGIHCOHINECEVEPCKNIGI 133
DB 129 GFCVINGSPCOHGCTCVDEGRASHA---SCLCPFGSGNFCETIVANCTPNPCENGV 184
OY 134 CTDLVANSCSCPEEFGRNC-----OYKC-----SGPLG 163
DB 185 CSDIGDPRCRCPAGFIDKTCSPVNCASSPCQNGTCLQHNVYBELCKPFTGTJC 244
OY 164 IEGGLISMOIT 175
DB 245 VKKRLSPQOYT 256

```

RESULT 11
 DTK_MOUSE STANDARD; PRT; 385 AA.
 AC 009163; 007645; 062206;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

```

DE DELTA-LIKE PROTEIN PRECURSOR (DLK) (PRADIPOCYTE FACTOR 1) (PREF-1)
DE (ADIPOCYTE DIFFERENTIATION INHIBITOR PROTEIN) [CONTAINS: FETAL ANTIGEN
DE 1 (FAL)].
GN DLK OR DLK OR PREF1 OR SCP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI-TaxId:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SWISS; TISSUE=Fibroblast;
RX MEDLINE-93179372; PubMed-8095043;
RA Laborda J., Sausville E.A., Hoffman T., Notario V.;
RT "dlk, a putative mammalian homeotic gene differentially expressed in
RT small cell lung carcinoma and neuroendocrine tumor cell line.";
RT J. Biol. Chem. 268:3817-3820(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-93272313; PubMed-8500166;
RA Smaas C.M., Sul H.S.;
RT "Pref-1, a protein containing EGF-like repeats, inhibits adipocyte
RT differentiation.";
RT Cell 73:725-734(1993).
RN [3]
RP SEQUENCE FROM N.A. AND VARIANT LYS-347 DEL.
RC TISSUE=Adrenal gland, and Placenta;
RX MEDLINE-95226449; PubMed-7711066;
RA Lee Y.L., Helman L., Hoffman T., Laborda J.;
RT "dlk, pg2 and Pref-1 mRNAs encode similar proteins belonging to the
RT EGF-like superfamily. Identification of polymorphic variants of this
RT RNA.";
RT Biochim. Biophys. Acta 1261:223-232(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA Maryama K., Nishijima S., Kuromitsu S., Ichikawa A., Masuda E.,
RA Takemoto T., Kohana H., Kawashima H.;
RL Submitted (AUG-1993) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-8 FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE-94325292; PubMed-7519443;
RA Smaas C.M., Green D., Sul H.S.;
RT "Structural characterization and alternate splicing of the gene
RT encoding the preadipocyte EGF-like protein pref-1.";
RL Biochemistry 33:9257-9265(1994).
RN [6]
RX CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE-97234561; PubMed-9118998;
RA Krogh T.N., Bachmann E., Teisner B., Skjoed K., Hoegjrup P.;
RT "Glycosylation analysis and protein structure determination of murine
RT fetal antigen 1 (mFAL) -- the circulating gene product of the
RT delta-like protein (dlk), preadipocyte factor 1 (Pref-1) and
RT stromal-cell-derived protein 1 (SCP-1) cDNAs.";
RL Eur. J. Biochem. 244:334-342(1997).
CC -1- FUNCTION: MAY HAVE A ROLE IN NEUROENDOCRINE DIFFERENTIATION.
CC INHIBITS ADIPOCYTE DIFFERENTIATION.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: A (SHOWN HERE), B, C,
CC C2, D AND D2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PRTM: N- AND O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: 212171; CAA78162.1;
CC EMBL: 015980; AAB60495.1;
CC EMBL: L12721; AAA37175.1;

```

DR	EMBL:	S71340:	-	NOT_ANNOTATED_CDS.
DR	EMBL:	D16847:	BAA04121.1:	-
DR	RSSP:	P00740:	11RX.	-
DR	GlycoSiteDB:	009163:	-	-
DR	MdD:	MdI:94900:	D1K1.	-
DR	InterPro:	IPR000152:	-	-
DR	InterPro:	IPR000561:	-	-
DR	Pfam:	PF00008:	Asx_6.	-
DR	PROSITE:	PS00010:	ASX_HYDROXYL; 1.	-
DR	PROSITE:	PS00022:	Egf_1; 5.	-
DR	PROSITE:	PS01186:	Egf_2; 6.	-
KW	Antigen:	Transmembrane;	EGF-like domain; Repeat; Glycoprotein; Signal.	-
KV	Alternative splicing:			-
FT	SIGNAL	1	23	BY SIMILARITY.
FT	CHAIN	24	385	DELTA-LIKE PROTEIN.
FT	CHAIN	24	305	FETAL ANTIGEN 1.
FT	DOMAIN	24	305	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	306	329	POTENTIAL.
FT	DOMAIN	330	385	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	24	55	EGF-LIKE 1.
FT	DOMAIN	33	86	EGF-LIKE 2.
FT	DOMAIN	88	125	EGF-LIKE 3.
FT	DOMAIN	127	168	EGF-LIKE 4.
FT	DOMAIN	172	208	EGF-LIKE 5.
FT	DOMAIN	210	247	EGF-LIKE 6.
FT	DISULFID	26	37	BY SIMILARITY.
FT	DISULFID	30	43	BY SIMILARITY.
FT	DISULFID	45	54	BY SIMILARITY.
FT	DISULFID	57	68	BY SIMILARITY.
FT	DISULFID	63	74	BY SIMILARITY.
FT	DISULFID	76	85	BY SIMILARITY.
FT	DISULFID	92	103	BY SIMILARITY.
FT	DISULFID	97	113	BY SIMILARITY.
FT	DISULFID	115	124	BY SIMILARITY.
FT	DISULFID	131	144	BY SIMILARITY.
FT	DISULFID	138	156	BY SIMILARITY.
FT	DISULFID	158	167	BY SIMILARITY.
FT	DISULFID	176	187	BY SIMILARITY.
FT	DISULFID	181	196	BY SIMILARITY.
FT	DISULFID	198	207	BY SIMILARITY.
FT	DISULFID	214	225	BY SIMILARITY.
FT	DISULFID	219	235	BY SIMILARITY.
FT	DISULFID	237	246	BY SIMILARITY.
FT	CARBOHYD	94	94	O-LINKED.
FT	CARBOHYD	100	100	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	165	165	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	174	174	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	216	216	O-LINKED.
FT	CARBOHYD	224	224	O-LINKED.
FT	CARBOHYD	258	258	O-LINKED.
FT	CARBOHYD	267	267	O-LINKED.
FT	CARBOHYD	271	271	O-LINKED (PARTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLC	211	303	MISSING (IN ISOPORM D).
FT	VARSPLC	211	303	MISSING (IN ISOPORM D2).
FT	VARSPLC	231	303	MISSING (IN ISOPORM B).
FT	VARSPLC	231	303	MISSING (IN ISOPORM C).
FT	VARSPLC	231	305	MISSING (IN ISOPORM C2).
FT	VARIANT	347	347	MISSING
FT	CONFLICT	250	250	R P (IN REF. 2).
FT	CONFLICT	320	385	LAVGATVAIVLNCKEWTYSNLRNHFFKKKKULLQNSGE LVANIFPEKIDDKTKKGDEEL -> CWAMPSESTSI AFRCCTCAITTCARRRRISCSITRAWSNRSSISSPRRL (IN REF. 2). TF -> ML (IN REF. 4). CRC64: E79664FEBA7A4FEFL
SEQ	SEQUENCE	385 AA:	41320 NM:	

	Query Match	Similarity	Best Local Match	Matches	Conservative	Indels	Gaps
0y	26	CDBNPOENGKICLGLAVGFSFCEDPGFTDPCNSVVEASDDEBPTSGAGCTTN-PC	83	9.1%	Pred. 244.5	DB 1	Length 385
				31.6%	Pred. 1.1e-11		
				50	18	Mismatches 65	

Db	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Db	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Db	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Db	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Db	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
Db	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80																				

Query Match
Best Local Similarity 27.58; Pred. No. 47e-07; Length 379;
Matches 55; Conservative 16; Mismatches 72; Indels 57; Gaps 10;

FT DISULFID 262 264 POTENTIAL.
FT DISULFID 273 282 POTENTIAL.
FT DISULFID 278 288 POTENTIAL.
FT DISULFID 294 296 POTENTIAL.
FT DISULFID 305 314 POTENTIAL.
FT DISULFID 310 320 POTENTIAL.
FT DISULFID 326 328 POTENTIAL.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 379 AA: 41590 MW: E3765F264B2BC9A CRC64;

DB 183 PNCENGICLPGLAVSFCSCPCDGTDPNCSSVEVASDEEFTSAGP---CTP--- 80
OY 81 -----NPGHNGTCEISEAYRDNTEIGVCKCPRGENGHICQHINECEVEPCKN 130
DB 238 GVNCDKANCSTCFNGGTC-----FYPCKICCPGLEGEQCE--LSKCP-QPCRN 284
OY 131 GIGCTDLVANYSCPCPGFMRNCOYK-CSGPLGIEG-----GIISNOOI 174
DB 285 GSKC---IGSKCKCPKGYGDLCKRPVCPGCGAHGTCHEPNKQCGREGHGNCKRY 341
OY 175 TASSTR---ALFGLQKWP 191
DB 342 GASIMAHAPRAGAGLERHTP 361

RESULT 13
WIFI_HUMAN STANDARD: PRT: 379 AA.
AC Q919M5;
DT 01-OCT-2000 (rel. 40, Created)
DT 01-OCT-2000 (rel. 40, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE WNT INHIBITORY FACTOR 1 PRECURSOR (WIF-1).
GN WIF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=99215557; PubMed=10201374;
RA Hsieh J.-C., Kodjabachian L., Rebber M.L., Rattner A.,
RT Smallwood P.M., Samos C.H., Nusse R., David I.B., Nathans J.;
RT "A new secreted protein that binds to Wnt proteins and inhibits their
activities.";
RL Nature 398:431-436(1999).
CC -1- FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY
CC -1- BE INVOLVED IN MESODERM SEGMENTATION.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF129222; AAD25402.1; -
CC
CC MIM: 605186; -
CC HSP: P00743; ICGF.
CC InterPro: IPR000561; -
CC InterPro: IPR002049; -
CC Pfam: PF00008; EGF; 5.

DR PRINTS; PRO0011; EGFAMININ.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 4.
KW Repeat; EGF-like domain; Signal; Developmental protein.
FT SIGNAL 1 28
FT DOMAIN 177 208 EGF-LIKE 1.
FT DOMAIN 209 240 EGF-LIKE 2.
FT DOMAIN 241 272 EGF-LIKE 3.
FT DOMAIN 273 304 EGF-LIKE 4.
FT DOMAIN 305 336 EGF-LIKE 5.
FT DISULFID 177 186 POTENTIAL.
FT DISULFID 182 192 POTENTIAL.
FT DISULFID 198 200 POTENTIAL.
FT DISULFID 209 218 POTENTIAL.
FT DISULFID 219 228 POTENTIAL.
FT DISULFID 230 232 POTENTIAL.
FT DISULFID 241 250 POTENTIAL.
FT DISULFID 246 256 POTENTIAL.
FT DISULFID 262 264 POTENTIAL.
FT DISULFID 273 282 POTENTIAL.
FT DISULFID 294 296 POTENTIAL.
FT DISULFID 305 314 POTENTIAL.
FT DISULFID 310 320 POTENTIAL.
FT DISULFID 326 328 POTENTIAL.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 379 AA: 41512 MW: 27782370A266784 CRC64;

Query Match
Best Local Similarity 25.68; Pred. No. 8.7e-07; Length 379;
Matches 51; Conservative 16; Mismatches 52; Indels 80; Gaps 10;

OY 28 PNCENGICLPGLAVSFCSCPCDGTDPNCSSVEVASDEEFTSAGPCTPMPCNG 87
DB 183 PGCGRMGFCNERRV-----CECPDGFYGHCEKALCIPRCNMGCLVTGFCICPPGFY 237
OY 88 TCEISEAYRDNTEIGVCKCPRGENGHIC-----Q 117
DB 223 LC-VTPGF-----CICPGRGYVNCDCANCSTCFNGSTCFPGKICPPGLEGD 272
OY 118 HINICEVEPCNKGIGCTDLVANYSCPCPGFMRNCOYK-CSGPLGIEG----- 166
DB 273 CEISKCP-QPCRNQKGC---IGSKCKCKSKGYGDLCKRPVCPGCGAHGTCHEPNKQCG 328
OY 167 -----GIISNOOTASTR 180
DB 329 QEGWHGHNCKRYEASLH 347

RESULT 14
FBI5_HUMAN STANDARD: PRT: 448 AA.
AC Q90BA5; 075966;
DT 01-OCT-2000 (rel. 40, Created)
DT 01-OCT-2000 (rel. 40, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE FIBULIN-5 PRECURSOR (FIBL-5) (DEVELOPMENTAL ARTERIES AND NEURAL CREST
DE EGF-LIKE PROTEIN) (DANCE) (URINE P50 PROTEIN) (UP50).
GN FBI5 OR DANCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Kostka G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
SQ SEQUENCE FROM N.A.

RX MEDLINE-99357779; PubMed-10428823;
 RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
 RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
 RA Matsumori A., Sasayama S., Chien K.R., Honjo T.,
 RT "DANCE, a novel secreted RGD protein expressed in developing,
 RT atherosclerotic, and balloon-injured arteries.";
 RL J. Biol. Chem. 274:22476-22483(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Urine;
 RA Zemel R., Sholto O., Shaul Y.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
 CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
 CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
 CC DEVELOPMENT AND REMODELING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEART, OVARY, AND
 CC COLON BUT ALSO IN KIDNEY, PANCREAS, TESTIS, LUNG, AND PLACENTA.
 CC NOT DETECTABLE IN BRAIN, LIVER, THYMUS, PROSTATE, OR PERIPHERAL
 CC BLOOD LEUCOCYTES.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AJ133490; CAB38568.1;
 DR EMBL: AF112152; AAD41767.1;
 DR EMBL: AF093118; AAC62107.1;
 DR MIM: 604580;
 DR InterPro: IPR000152;
 DR InterPro: IPR000561;
 DR InterPro: IPR001881;
 DR Pfam: PF00008; EGF_4;
 DR PROSITE: PS00010; ASX_HYDROXYL_4;
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 4;
 DR PROSITE: PS01187; EGF_CA; 6;
 KM Repeat: EGF-like domain; Calcium-binding; Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 448
 FT DOMAIN 24 69 EGF-LIKE 1, DIVERGENT.
 FT DOMAIN 127 167 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 207 246 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 131 144 BY SIMILARITY.
 FT DISULFID 138 153 BY SIMILARITY.
 FT DISULFID 155 166 BY SIMILARITY.
 FT DISULFID 172 181 BY SIMILARITY.
 FT DISULFID 177 190 BY SIMILARITY.
 FT DISULFID 192 205 BY SIMILARITY.
 FT DISULFID 211 220 BY SIMILARITY.
 FT DISULFID 217 230 BY SIMILARITY.
 FT DISULFID 232 245 BY SIMILARITY.
 FT DISULFID 251 262 BY SIMILARITY.
 FT DISULFID 258 271 BY SIMILARITY.
 FT DISULFID 273 286 BY SIMILARITY.
 FT DISULFID 292 305 BY SIMILARITY.
 FT DISULFID 299 314 BY SIMILARITY.
 FT DISULFID 320 332 BY SIMILARITY.
 FT CARBOHYD 283 283 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 296 296 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 69 70 IP -> HS (IN REF. 3).
 FT CONFLICT 147 148 TE -> MK (IN REF. 3).
 SQ SEQUENCE 448 AA, 50180 MW, 19FCA51FDA328003 CRC64;

Query Match 5.6%; Score 150; DB 1; Length 448;
 Best Local Similarity 25.0%; Pred. No. 0.00028;
 Matches 49; Conservative 24; Mismatches 51; Indels 72; Gaps 12;
 QY 8 WLVLGSLGVPOFGKGDICDPNPEANGICLPGLAVGFSFSCPCPGFT---DNGCSYV 63
 DB 160 WLEGGQCLDI-----DECRVGYCOOLCANVP---GSYCTGCPFTLNDGNSCODVN 209
 QY 64 EVASDEEPTSAQCTPPNCPHNGCTCEISEAYRGDTFICVCKPGRF---NGHCQHN 119
 DB 210 ECA-----TENPCVQ--TCV-----NTYGSFICRCPGYELBEDGVCS-D 247
 QY 120 INCEVEP--CKNGICGDIIVANVSCPCGPEF---GNQC----- 155
 DB 248 WDCRSSEFLQHE--CVNQGTYFCSCPPQITLLDNRSCODINCEHRHNTCNLDQTC 305
 QY 156 -----YKCSGFLGIE 165
 DB 306 YNLGGFKCIDPIRCE 321
 RESULT 15
 FBL5_MOUSE
 ID FBL5_MOUSE STANDARD: PRT; 448 AA.
 AC QPWH9;
 DT 01-OCT-2000 (rel. 40; Created)
 DT 01-OCT-2000 (rel. 40; Last sequence update)
 DT 01-OCT-2000 (rel. 40; Last annotation update)
 DE FIBULIN-5 PRECURSOR (FBL-5) (DEVELOPMENTAL ARTERIES AND NEURAL CREST
 DE EGF-LIKE PROTEIN) (DANCE).
 GN FBLN5 OR DANCE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99357779; PubMed-10428823;
 RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
 RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
 RA Matsumori A., Sasayama S., Chien K.R., Honjo T.,
 RT "DANCE, a novel secreted RGD protein expressed in developing,
 RT atherosclerotic, and balloon-injured arteries.";
 RL J. Biol. Chem. 274:22476-22483(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Urine;
 RA Zemel R., Sholto O., Shaul Y.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
 CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
 CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
 CC DEVELOPMENT AND REMODELING.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEART, OVARY, AND
 CC COLON BUT ALSO IN KIDNEY, PANCREAS, TESTIS, LUNG, AND PLACENTA.
 CC NOT DETECTABLE IN BRAIN, LIVER, THYMUS, PROSTATE, OR PERIPHERAL
 CC BLOOD LEUCOCYTES.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF112151; AAD41767.1;
 DR MIM: 604580;
 DR InterPro: IPR000152;
 DR InterPro: IPR000561;
 DR InterPro: IPR001881;
 DR Pfam: PF00008; EGF_4;
 DR PROSITE: PS00010; ASX_HYDROXYL_4;
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 4;
 DR PROSITE: PS01187; EGF_CA; 6;
 KM Repeat: EGF-like domain; Calcium-binding; Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 448
 FT DOMAIN 24 69 EGF-LIKE 1, DIVERGENT.
 FT DOMAIN 127 167 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 207 246 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 131 144 BY SIMILARITY.
 FT DISULFID 138 153 BY SIMILARITY.
 FT DISULFID 155 166 BY SIMILARITY.
 FT DISULFID 172 181 BY SIMILARITY.
 FT DISULFID 177 190 BY SIMILARITY.
 FT DISULFID 192 205 BY SIMILARITY.
 FT DISULFID 211 220 BY SIMILARITY.
 FT DISULFID 217 230 BY SIMILARITY.
 FT DISULFID 232 245 BY SIMILARITY.
 FT DISULFID 251 262 BY SIMILARITY.
 FT DISULFID 258 271 BY SIMILARITY.
 FT DISULFID 273 286 BY SIMILARITY.
 FT DISULFID 292 305 BY SIMILARITY.
 FT DISULFID 299 314 BY SIMILARITY.
 FT DISULFID 320 332 BY SIMILARITY.
 FT CARBOHYD 283 283 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 296 296 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 69 70 IP -> HS (IN REF. 3).
 FT CONFLICT 147 148 TE -> MK (IN REF. 3).
 SQ SEQUENCE 448 AA, 50180 MW, 19FCA51FDA328003 CRC64;

```
FT CHAIN 24 448 FIBULIN-5, DIVERGENT.
FT DOMAIN 24 69 EGF-LIKE 1, DIVERGENT.
FT DOMAIN 127 167 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 246 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 153 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 258 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 448 AA; 50193 MW; F15CC70CCFBFC97 CRC64;
```

Query Match 5.6%; Score 149; DB 1; Length 448;
Best Local Similarity 25.4%; Pred. No. 0.00033;
Matches 50; Conservative 50; Mismatches 49; Indels 74; Gaps 13;

```
QY 8 WLVLGSLGVPOFGKGDICDPNCPENGICLPGLAVGSFSCPCPDGFT---DPNCSSVY 63
DB 160 WLEGGQLDI-----DECRIGYCOQLCANP---GSYSTCNPEFTLNDGRSCQV 209
QY 64 EVASDEEPTSAGPC-TPNCPHNGTGCEISEAYRGDTFGYVCKCPRGF---NSIHQOH 118
DB 210 E-----CETENPCVO--TCV-----NTYGSFTICRCDPGYELEDGDIHCS- 246
QY 119 NINECEVER--CKNGICTDLVANYSCPCPEFM---GRNCO----- 155
DB 247 DMDECSFSEFLQHE--CVNQPGSYFCSCPFGYVLLDNNRSCODINECEHNRHTCTSLQT 304
QY 156 -----YKCSGPICIE 165
DB 305 CYNIGGSFKCIDPISCE 321
```

Search completed: May 23, 2001, 08:40:07
Job time: 118 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:38:29 ; Search time 31.5 Seconds
(Without alignments)
1789.746 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513
Perfect score: 2675
Sequence: 1 MKRSVAVMLVGLSLGVPQF.....MSWGRITLASLLGCTEEE 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 segs, 117207915 residues

Total number of hits satisfying chosen parameters: 310160

Minimum DB seq length: 0
Maximum DB seq length: 481

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTRMBL_15: *
2: sp-archaea: *
3: sp-bacteria: *
4: sp-fungi: *
5: sp-human: *
6: sp-invertebrate: *
7: sp-mammal: *
8: sp-mmc: *
9: sp-organelle: *
10: sp-phage: *
11: sp-plant: *
12: sp-rodent: *
13: sp-vertebrate: *
14: sp-virus: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	263.5	98.4	4 043854	043854 homo sapien
2	2529.5	94.6	11 033474	033474 mus musculu
3	1177	44.0	11 091X9	091X9 mus musculu
4	1154.5	43.2	11 09WTS3	09WTS3 mus musculu
5	972	36.3	363 6 077718	077718 equus caball
6	464.5	17.4	216 4 014286	014286 homo sapien
7	428.5	16.0	192 11 09QW01	09QW01 mus musculu
8	318	11.9	406 5 025059	025059 heliocladi
9	286.5	10.7	156 5 026661	026661 strongyloce
10	279	10.4	224 11 09R1M6	09R1M6 mus musculu
11	272	10.2	263 4 099734	099734 homo sapien
12	269.5	10.1	263 4 099740	099740 homo sapien
13	268.5	10.0	434 11 055139	055139 mus musculu
14	259.5	9.7	192 4 09UPM2	09UPM2 homo sapien
15	257	9.6	305 6 09N028	09N028 macaca fasc
16	253.5	9.5	308 6 046370	046370 bos taurus
17	250.5	9.4	383 11 070534	070534 rattus norv
18	250.5	9.4	383 11 062779	062779 rattus norv
19	239	8.9	387 11 006007	006007 mus musculu

20	205	7.7	78	4	P78328	P78328 homo sapien
21	204.5	7.6	252	4	Q9NTR1	Q9NTR1 homo sapien
22	185.5	6.9	379	11	Q9W0A1	Q9W0A1 mus musculu
23	184.5	6.9	355	4	Q9Y506	Q9Y506 homo sapien
24	183	6.8	374	13	Q9W6F8	Q9W6F8 xenopus lae
25	182.5	6.8	315	4	Q9UPK6	Q9UPK6 homo sapien
26	182	6.8	379	4	Q9Y5W5	Q9Y5W5 homo sapien
27	172	6.4	378	13	Q9W6F9	Q9W6F9 brachydactilo
28	170	6.4	220	11	Q63404	Q63404 rattus norv
29	169	6.3	420	5	Q18375	Q18375 caenorhabd
30	168.5	6.3	372	5	Q21756	Q21756 caenorhabd
31	164.5	6.1	445	5	Q9W3W5	Q9W3W5 drosophila
32	159	5.9	176	13	Q9PUC7	Q9PUC7 gallus gall
33	159	5.9	349	5	Q17377	Q17377 caenorhabd
34	158.5	5.9	379	5	Q19922	Q19922 caenorhabd
35	154.5	5.8	287	4	Q76101	Q76101 homo sapien
36	152.5	5.7	255	11	Q08745	Q08745 mus musculu
37	152.5	5.7	304	11	Q08744	Q08744 mus musculu
38	152	5.7	434	5	Q9W0A0	Q9W0A0 drosophila
39	147	5.5	259	5	Q93519	Q93519 caenorhabd
40	146	5.5	323	5	Q9W6G4	Q9W6G4 drosophila
41	144	5.4	247	5	Q9W6F8	Q9W6F8 drosophila
42	142	5.3	78	4	Q9U119	Q9U119 homo sapien
43	140	5.2	144	13	Q12973	Q12973 gallus gall
44	138.5	5.2	229	11	Q9QY21	Q9QY21 mus musculu
45	138.5	5.2	266	11	Q9R1K1	Q9R1K1 rattus norv

ALIGNMENTS

RESULT 1	
ID 043854	PRELIMINARY; PRT: 480 AA.
AC 043854; 043855;	
DT 01-JUN-1998 (TREMBLrel. 06, Created)	
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)	
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)	
DE INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.	
GN DEL1.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OC NCBI_TaxID=9606;	
RN 11)	
RC SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.	
RC TISSUE=EMBRYONIC LUNG;	
RX MEDLINE=98083109; PubMed=9420328;	
RA Hidaï C., Zupancic T.J., Penta K., Mikhail A., Kawana M.,	
RA Quertermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,	
RA Auerbach R., Hogan B.L.M., Snodgrass R., Quertermous T.;	
RT "Cloning and characterization of developmental locus-1: an	
RT embryonic endothelial cell protein that binds the alphavbeta3 integrin	
RT receptor".	
RT Genes Dev. 12:21-33(1998).	
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH	
CC INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS	
CC FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN	
CC REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC	
CC DEVELOPMENT.	
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.	
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND	
CC SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.	
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.	
CC EMBL: U70312; AAC02648.1; -	
DR EMBL: U70313; AAC02649.1; -	
DR HSSP: P00740; I1XA.	
DR INTERPRO: IPR0000152; -	
DR INTERPRO: IPR000421; -	
DR INTERPRO: IPR000561; -	
DR INTERPRO: IPR000742; -	
DR INTERPRO: IPR001881; -	

FT DISULFID 107 116 BY SIMILARITY.
 FT DISULFID 158 314 BY SIMILARITY.
 FT DISULFID 301 305 BY SIMILARITY.
 FT DISULFID 319 476 BY SIMILARITY.
 FT VARSPLIC 218 221 INDO -> VIVG (IN SHORT ISOFORM).
 FT VARSPLIC 222 480 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 480 AA; 53740 MW; 4CD91EF9261714D CRC64;

Query Match 94.6%; Score 2529.5; DB 11; Length 480;
 Best Local Similarity 94.4%; Pred. No. 6,4e-219;
 Matches 454; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MKRSYAVWLVLGLSLGVPQKGDICDPNCEGICLPOLAVGSFSCPCPDGTPDCS 60
 1 MHILVAALVLGLSLGVPQKGDICDPNCEGICLPOLAVGSFSCPCPDGTPDCS 60
 DB 1 MHILVAALVLGLSLGVPQKGDICDPNCEGICLPOLAVGSFSCPCPDGTPDCS 60
 QY 61 SVVEVASEDEEPTPSAGCPNCHNGNCEISEAYRGDTFTGYVCKCPGFGHICQHN 120
 61 SVVEVASEDEEPTPSAGCPNCHNGNCEISEAYRGDTFTGYVCKCPGFGHICQHN 120
 DB 61 SVVEVASEDEEPTPSAGCPNCHNGNCEISEAYRGDTFTGYVCKCPGFGHICQHN 120
 QY 121 NCEVEPCCKGICITDLVANYSCGCPGFGHICQHNCEISEAYRGDTFTGYVCKCPGFGHICQHN 180
 121 NCEVEPCCKGICITDLVANYSCGCPGFGHICQHNCEISEAYRGDTFTGYVCKCPGFGHICQHN 180
 DB 121 NCEVEPCCKGICITDLVANYSCGCPGFGHICQHNCEISEAYRGDTFTGYVCKCPGFGHICQHN 180
 QY 181 RALFSLQKWPYARLKKGLINAMTAANDRKRMVQINLQRRKRVGTITOGAKRIGS 240
 181 RALFSLQKWPYARLKKGLINAMTAANDRKRMVQINLQRRKRVGTITOGAKRIGS 240
 DB 181 RALFSLQKWPYARLKKGLINAMTAANDRKRMVQINLQRRKRVGTITOGAKRIGS 240
 QY 241 PEYIKFYKIASNDGKTMAVYKGTNDMVEFGNIDNNTPVANSFTPIKAQVRLYPQ 300
 241 PEYIKFYKIASNDGKTMAVYKGTNDMVEFGNIDNNTPVANSFTPIKAQVRLYPQ 300
 DB 241 PEYIKFYKIASNDGKTMAVYKGTNDMVEFGNIDNNTPVANSFTPIKAQVRLYPQ 300
 QY 301 VCRHCTLRMELLCGCELSGCEPLMGKSHIODYOITASSIFPTLMDMFTWEPKARLD 360
 301 VCRHCTLRMELLCGCELSGCEPLMGKSHIODYOITASSIFPTLMDMFTWEPKARLD 360
 DB 301 VCRHCTLRMELLCGCELSGCEPLMGKSHIODYOITASSIFPTLMDMFTWEPKARLD 360
 QY 361 KGGKVNATSGHNDOSOMLOVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMW 420
 361 KGGKVNATSGHNDOSOMLOVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMW 420
 DB 361 KGGKVNATSGHNDOSOMLOVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMW 420
 QY 421 TYVQDEKRRKRVQGNFNDHRRKNVIDPPIYARHRIIPMSWYGRITLASSELLGCEE 480
 421 TYVQDEKRRKRVQGNFNDHRRKNVIDPPIYARHRIIPMSWYGRITLASSELLGCEE 480
 DB 421 TYVQDEKRRKRVQGNFNDHRRKNVIDPPIYARHRIIPMSWYGRITLASSELLGCEE 480
 QY 481 E 481
 DB 480 E 480

RESULT 3

Q9RLX9

ID Q9RLX9

AC Q9RLX9

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, last sequence update)

DT 01-OCT-2000 (Tremblrel. 15, last annotation update)

DE MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 LONG FORM.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=BAIB/C; TISSUE=MAMMARY GLAND;

RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;

RT "Lactation-dependent expression of an mRNA splice variant with an exon

RT for a multiply O-glycosylated domain of mouse milk fat globule

RT glycoprotein MFG-E8."

RL Biochem. Biophys. Res. Commun. 254:522-528(1999).

DR EMBL; AB021130; BAA35180.1; -.

DR HSSP; P00740; 1IXA.

DR INTERPRO; IPR000421; -.

DR INTERPRO; IPR000561; -.
 DR INTERPRO; IPR001092; -.
 DR INTERPRO; IPR001438; -.
 DR EFAM; PF00008; EGF_2.
 DR EFAM; PF00734; EGF_2-type C; 2.
 DR PRINTS; PR00010; EGFBL00.
 DR PROSITE; PS00022; EGF_1; UNKNOWN.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01285; FAS5C_1; 2.
 DR PROSITE; PS01286; FAS5C_2; 2.
 SQ SEQUENCE 463 AA; 51269 MW; D719D2BE090E6427 CRC64;

Query Match 44.0%; Score 1177; DB 11; Length 463;
 Best Local Similarity 49.1%; Pred. No. 1.7e-97;
 Matches 225; Conservative 72; Mismatches 139; Indels 22; Gaps 8;

QY 23 GICDPNCEGICITPLGLAVGSFSCPCPDGTPDCSSVVEVASEDEEPTPSAGCPN 82
 23 GICDPNCEGICITPLGLAVGSFSCPCPDGTPDCSSVVEVASEDEEPTPSAGCPN 82
 DB 23 GICDPNCEGICITPLGLAVGSFSCPCPDGTPDCSSVVEVASEDEEPTPSAGCPN 82
 QY 83 CHNGTCEIS-EARSGDTFTGYVCKCPGFGHICQHNCEISEAYRGDTFTGYVCKCPGFGHICQHN 141
 83 CHNGTCEIS-EARSGDTFTGYVCKCPGFGHICQHNCEISEAYRGDTFTGYVCKCPGFGHICQHN 141
 DB 83 CHNGTCEIS-EARSGDTFTGYVCKCPGFGHICQHNCEISEAYRGDTFTGYVCKCPGFGHICQHN 141
 QY 142 SCECPGEF--MGRNCOYKSGPLGIEGIIISNOQTASSSTRALFSLQKWPYARLKK 199
 142 SCECPGEF--MGRNCOYKSGPLGIEGIIISNOQTASSSTRALFSLQKWPYARLKK 199
 DB 142 SCECPGEF--MGRNCOYKSGPLGIEGIIISNOQTASSSTRALFSLQKWPYARLKK 199
 QY 200 GLINAMTAANDRKRMVQINLQRRKRVGTITOGAKRIGSPYIRFYKIASNDGKTMA 259
 200 GLINAMTAANDRKRMVQINLQRRKRVGTITOGAKRIGSPYIRFYKIASNDGKTMA 259
 DB 200 GLINAMTAANDRKRMVQINLQRRKRVGTITOGAKRIGSPYIRFYKIASNDGKTMA 259
 QY 249 FQDE-SGGRKEPLGNDNNSLKVNNENPTLEAOYRLPYSCRGCTLRPELLECELHG 307
 249 FQDE-SGGRKEPLGNDNNSLKVNNENPTLEAOYRLPYSCRGCTLRPELLECELHG 307
 DB 249 FQDE-SGGRKEPLGNDNNSLKVNNENPTLEAOYRLPYSCRGCTLRPELLECELHG 307
 QY 320 GCEPLMGKSHIODYOITASSIFPTLMDMFTWEPKARLDKGGKVNATSGHNDOSOMLO 379
 320 GCEPLMGKSHIODYOITASSIFPTLMDMFTWEPKARLDKGGKVNATSGHNDOSOMLO 379
 DB 320 GCEPLMGKSHIODYOITASSIFPTLMDMFTWEPKARLDKGGKVNATSGHNDOSOMLO 379
 QY 361 KGGKVNATSGHNDOSOMLOVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMW 420
 361 KGGKVNATSGHNDOSOMLOVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMW 420
 DB 361 KGGKVNATSGHNDOSOMLOVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMW 420
 QY 380 OVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMWYVQDEKRRKRVQGNFND 439
 380 OVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMWYVQDEKRRKRVQGNFND 439
 DB 380 OVDLVPYKGTITOGAKDGHVQFVSGYKLANSDGEMWYVQDEKRRKRVQGNFND 439
 QY 440 NDHRRKNVIDPPIYARHRIIPMSWYGRITLASSELLGCEE 477
 440 NDHRRKNVIDPPIYARHRIIPMSWYGRITLASSELLGCEE 477
 DB 440 NDHRRKNVIDPPIYARHRIIPMSWYGRITLASSELLGCEE 477

RESULT 4

Q9WTS3

ID Q9WTS3

AC Q9WTS3

DT 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, last sequence update)

DT 01-OCT-2000 (Tremblrel. 15, last annotation update)

DE MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 S.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=BAIB/C; TISSUE=MAMMARY GLAND;

RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;

RT "Lactation-dependent expression of an mRNA splice variant with an exon

RT for a multiply O-glycosylated domain of mouse milk fat globule

RT glycoprotein MFG-E8."

RL Biochem. Biophys. Res. Commun. 254:522-528(1999).

DR EMBL; AB025280; BAA76386.1; -.

DR HSSP; P00740; 1IXA.

```

DR INTERPRO: IP000421; -
DR INTERPRO: IP000561; -
DR INTERPRO: IP001092; -
DR INTERPRO: IP001438; -
DR PFAM: PF00008; EGF_2;
DR PFAM: PF00754; F5_E8_Type-C; 2.
DR PRINTS: PR00010; EGRBLOOD.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FAS3C_1; 2.
DR PROSITE: PS01286; FAS3C_2; 2.
DR SEQSOURCE 426 AA; 47197 MW; B182FA20629881A CRC64;

```

Query Match	Local Similarity	43.2%	Score 1154.5	DB 11	Length 426
Best Local Similarity	48.0%	Pred. No. 1.6e-95			
Matches 219	Conservative 66	Mismatches 116	Indels 55	Gaps 7	
OY	23	GDIDPDPNCPNGGICLPGLAVSGFSFCSCDDGTPTPNCSSVAVASDEEPTSGACGCPNP	82		
Db	25	GPDCOSSICLNGGCTCLNG-QDNDIYCLDCEPFTGACNCE-----TERGCSGNP	72		
OY	83	CHAGGCGEIS-EAARGPTFGYCKCKPREFNHHOHNIMECEVECKKNGICLTDIVANY	141		
Db	73	CYADAKCLVTLDTRGDIPTFETICQCPVYSGICHENG-----	110		
OY	142	SCCCPPEFMGNCQYCKSGPLGEGITISNOQIAPASSPHALGLOKWTPIYARANKGL	201		
Db	111	-----CSTOLMEGALIDSDISMSVYMGFMGLDROMPELARIYRTGI	154		
OY	202	INANTALENDMKRWIOINLQKRMATPTOYITOGAKRTIGSEPIYIKFKIAYSDNGKITMAY	261		
Db	155	VANATASNTDS-KETIDYVNLNRKARSGVMTTQGSKRGRALERTKTKVAYSLDGKKEPTI	213		
OY	262	KYKGINEDMYRGKINDVNTPYANSFPFKPAQYARLTPQYCRNCHLRLMELGELGSCS	321		
Db	214	QDE-SGGDEFGMLDNNLSKVMNMFILIAQYIDLTPVSGHGGCTLRLLELGLCGLHS	272		
OY	322	EPLAKMSGHIODYITLSSIFRLIAMDFTWYEPKRALDKQGVANVATSGHDSOWILOV	381		
Db	273	EPGLGLNNMTIPDSQMSASSSTKTKWNLRAKAGMTPHILGRIDMGQKINMTATQSSAKENWY	332		
OY	382	DLLVLPKYTGITITOGADGCHVDFVGSYKLAISNDGEHMTVYDEKQRKDKVFOGENDD	441		
Db	333	DLTGQKQVYGIITOGADGCHIDHIDYAVASVYVHSDSGVQWMTYV--EPQSSKVFQGLDNN	390		
OY	442	THRKAVIDPPIYARHRIILPWSYNGIILASELGC	477		
Db	391	SHKKNIFKEPMAHYVAVLPVSNHNNITLRLLELGC	426		

RESULT	5			
077718				
ID	077718	PRELIMINARY:	PRT:	363 AA.
AC	077718;			
DT	01-NOV-1998	(TRMB rel. 08, Created)		
DT	01-NOV-1998	(TRMB rel. 08, Last sequence update)		
DT	01-MAY-2000	(TRMB rel. 13, Last annotation update)		
DE	SHEER-MEMBRANE ASSOCIATED PROTEIN P47 (FRAGMENT).			
OS	Equus caballus (Horse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.			
OX	NCBI_TaxId:9796;			
RN	11			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TESTIS;			
RA	Genetel M., Toepfer-Petersen E.;			
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.			
EMBL	AJ010121; CA090101.1; -			
DR	INTERPRO: IPR000421; -			
DR	INTERPRO: IPR000561; -			
DR	PF00008; EGF. 1.			
PPAM	PF00754; F5_P8_Type-C. 2.			

DR	PROSITE:	PS00022;	EGF_1;	UNKNOWN_2.
DR	PROSITE:	PS01186;	EGF_2;	2.
DR	PROSITE:	PS01285;	PA5BC_1;	1.
KW	Spem.			
FT	NON_TER	1	1	
FT	NON_TER	363	363	
SQ	SEQUENCE	363 AA;	40744 MM;	1F8B6395AF2338D CRC64;

Query Match	36.38;	Score 912;	DB 6;	Length 363;
Best Local Similarity	43.0%;	Pred. No. 3.2e-79;		
Matches 187; Conservative	64;	Mismatches 110;	Indels 74;	Gaps 6;

[illegible]

RESULT	6			
ID	014286	PRELIMINARY:	PRT:	216 AA.
AC	014286:			
DT	01-NOV-1996 (TREMblrel. 01, Created)			
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)			
DT	01-MAY-2000 (TREMblrel. 13, Last annotation update)			
DE	COAGULATION FACTOR VIII ASSOCIATED PROTEIN B.			
GN	F8B.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE:93052386; PubMed:1427887;			
RA	Levinson B., Kenrick S., Gamel P., Fisher K., Gleschler J.,			
RT	"Evidence for a third transcript from the human factor VIII gene."			
RL	Genomics 14:585-589(1992).			
EMBL	M90707; AAAS6466.1; -.			
DR	HSP: P00451; ICFG.			
DR	INTERPRO: IPR000421; -.			
DR	PRAM: PF00754; P5_P8_type_C; 1.			
DR	PROSITE: PS01285; FA58C_1; UNKNOWNM.1.			
DR	PROSITE: PS01286; FA58C_2; 2.			
QO	SEQUENCE 216 AA; 24641 MW; 6C82D4F89E35A376 CRC64;			

Query Match 17.4%; Score 464.5; DB 4; Length 216;
Best Local Similarity 46.7%; Pred. No. 66-34; Mismatches 75; Indels 7; Gaps 3;
Matches 99; Conservative 31;

QY 271 VFGNDNTNPTANSTPTPIAAYVRLPYOVCRRHCTLMELLCGELSGCEPLGKMSGH 330
DB 9 VFGNDSSGKHNIFNPITARIKHPHYSIRSTLMLMGDLNCSMPLEKMSKA 68
QY 331 IODYOTASSIFRTLMOMF-TWEPKARLDKOGKYNANTSGHNDQGLQVLDLVETKY 389
DB 69 ISDAOTASSYF---TMMFATWSPSKARLHLOGRSNAMPQVNNPKMLQVDTOKTKMY 124
QY 390 TGIITGAKDFHVOFGYSKLAAYNGEHWYVODEKORCKDPVQGFNDTHRKAVD 449
DB 125 TGVTTGVKSLTSMYKVEPLISSODGQHWTLF--FQKGYKVPQGNDSFTPVNSLD 182
QY 450 PPIYARHRLPMSWYGRITLASELCTEE 481
DB 183 PPIITRRLRHPSWYGRITLASELCTEAD 214

RESULT 7

Q9QWQ1 PRELIMINARY; PRT; 192 AA.
AC Q9QWQ1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE MILK FAT GLOBULE GLYCOPROTEIN (FRAGMENT).
MFC-E8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
RT Lactation-dependent expression of an mRNA splice variant with an exon
RT for a multiply O-glycosylated domain of mouse milk fat globule
RT glycoprotein MFC-E8.
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB021647; BAA36214.1;
DR EMBL; AB021647; BAA36214.1; JOINED.
DR INTERPRO; IPR000421;
DR INTERPRO; IPR000561;
DR INTERPRO; IPR001092;
DR PFAM; PF00008; EGF_1.
DR PFAM; PF00754; F5_F8_type_C; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01285; FA5RC_1; 1.
DR NON_TER 1
FT NON_TER 192
SQ SEQUENCE 192 AA; 21134 MW; E8110C03A89689F4 CRC64;

Query Match 16.0%; Score 428.5; DB 11; Length 192;
Best Local Similarity 45.4%; Pred. No. 8,7e-31;
Matches 83; Conservative 33; Mismatches 60; Indels 7; Gaps 4;

QY 77 PCTNPCHNGTCEIS-EAYRQDPTIGVCKPGRFGNHCNHNINEVEEPCNGKIGCT 135
DB 1 PCSNPICYNDACKLVTLDFORGDIFTEYLICCPGVSGIHCEFTENYVNLD---GEWYFT 57
QY 136 DLVANYSCGEGPGEF--MGRNCCYKCSGSLGEGGIINQOITASSYHRLFGLOKWPYY 193
DB 58 TAVNPTAVPTPAPTPDLNNLASRCSQGLGEGAIADQISASSVYMGFMGLQMGPEL 117
QY 194 ARUKKGGIINAAATENDRKRMIQINLRKMRVTGVTIGGAKRGSPYIKFYKIAVSN 253
DB 118 ARLRITGIVNAWTAISYDS-KPMIOVNLRLMKRYSVMTGGASRAGRATYKTFVVAISL 176

QY 254 DGR 256
DB 177 DGR 179

RESULT 8

Q25059 PRELIMINARY; PRT; 406 AA.
AC Q25059;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE FIBROPELIN III (FRAGMENT).
OS Helicidaris erythrogramma (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinometridae;
OC Helicidaris.
NCBI_TaxID=7634;
RN [1]
RP SEQUENCE FROM N.A.
RA Bisgrove B.W.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L3862; AAA28996.1;
DR HSSP; P00740; IEDM
DR INTERPRO; IPR000083;
DR INTERPRO; IPR000088;
DR INTERPRO; IPR000152;
DR INTERPRO; IPR000561;
DR INTERPRO; IPR000742;
DR INTERPRO; IPR001010;
DR INTERPRO; IPR001438;
DR INTERPRO; IPR001881;
DR INTERPRO; IPR001947;
DR INTERPRO; IPR002049;
DR PFAM; PF00008; EGF; 7.
DR PFAM; PF01382; Avidin; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGFBLMIN.
DR PRINTS; PR00012; ENTPEL.
DR PRINTS; PR00286; CHARYBDOTOXIN.
DR PRINTS; PR00287; THIONIN.
DR PRINTS; PR00709; AVIDIN.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_7.
DR PROSITE; PS00022; EGF_1; UNKNOWN_7.
DR PROSITE; PS00577; AVIDIN; 1.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS01187; EGF_CA; 5.
KW Glycoprotein; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 406 AA; 43475 MW; EC89CCE8A8B7E89A CRC64;

Query Match 11.9%; Score 318; DB 5; Length 406;
Best Local Similarity 33.9%; Pred. No. 2e-20;
Matches 59; Conservative 23; Mismatches 54; Indels 38; Gaps 4;

QY 23 GDICDPCNCGGICLGLAVSGSFSCGPDGFTDPCNSVVEVASD----- 68
DB 13 GDCNPNPCQNAACID--QYNDIEICPPTGDNCTDIDVCAAPCRNGACVQDVN 70
QY 69 -----EEPTSAEPTPNPCNGGICELSEAYRQDPTIGVCKPGRFGNHC 116
DB 71 GYTCNCIDGFDGDCENNINECASNPCQNGVCI-----DGVNGFVCTCGGTGTLG 123
QY 117 QHNIINEVEEPCNGKIGICTDLVANYSCGEGPGEFGRNCQ--YKSGPLGIEGG 167
DB 124 ETIDIDECASNPCQNGVCTDLVANYTCDCIAGTGSNCETININECASNPCLNG 177
RESULT 9
Q26661 PRELIMINARY; PRT; 156 AA.
AC Q26661;

```

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE EPIDERMAL GROWTH FACTOR (FRAGMENT).
GN EGF.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
CC NCBI_TaxID=7668;
RN
RP
RP SEQUENCE FROM N.A.
RX MEDLINE=90112459; PubMed=2514273;
RA Delgado-Illio-Reynoso M.G., Kollo D.R., Hursh D.A., Raff R.A.;
RT "Structural analysis of the uesg gene in the sea urchin
RT Strongylocentrotus purpuratus reveals more similarity to vertebrate
RT than to invertebrate genes with EGF-like repeats."
RL J. Mol. Evol. 29:314-327(1989).
DR EMBL: X17531; CAA3572.1; -.
DR HSSP: P07204; IEGD.
DR INTERPRO: IPR000088; -.
DR INTERPRO: IPR000152; -.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR000742; -.
DR INTERPRO: IPR001010; -.
DR INTERPRO: IPR001438; -.
DR INTERPRO: IPR001881; -.
DR INTERPRO: IPR001947; -.
DR PFAM: PF00008; EGF_3.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00286; CHARYD0XIN.
DR PRINTS: PR00287; THIONIN.
DR PRINTS: PR00709; AVIDIN.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
DR PROSITE: PS01186; EGF_2; UNKNOWN_3.
DR PROSITE: PS01187; EGF_CA; 2.
KW Glycoprotein; EGF-like domain.
FT
FT NON_TER 1
FT TER 156
FT SEQUENCE 156 AA; 16485 MW; E64360DBB029B1A2 CRC64;

Query Match 10.7%; Score 286.5; DB 5; Length 156;
Best Local Similarity 36.7%; Pred. No. 3.7e-18;
Matches 58; Conservative 20; Mismatches 55; Indels 25; Gaps 5;

QY 24 DIDDPPNCGGICPGIAVGSFCPCPGFTDPKSSVYVVASDEEPTASGCTPAPC 83
DB 2 DECASMPCLNGSACIE--MNGYTCCVAGTYGVC-----TDIDECASAPC 47
QY 84 HNGGCEISEARQDPTFGVYCKPFGFNGICQHNINECEVEPCCKNGICTDLVANTSC 143
DB 48 QNGSVGT-----DTINGIACVPGFTGSCFETNINECASLPCQNGGLCIDGIAGTYC 100
QY 144 ECPGEMRNCYKCSGFLGEGGIIIS--NQGITASSST 179
DB 101 QCRLLGYIGVCEE--VSCDLEGMVMECDVITTKT 136

RESULT 10
Q9R1M6 PRELIMINARY; PRT; 224 AA.
AC Q9R1M6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)
DE X-LINKED JUVENILE RETINOSCHISIS PROTEIN.
GN XLR51.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;

```

```

RN
RN SEQUENCE FROM N.A.
RA Reid S.N.M., Akhmedov N.B., Parlev N.I., Kozak C.A., Danciger M.,
RA Farber D.B.;
RT "The mouse X-linked juvenile retinoschisis cDNA: expression in
RT photoreceptors."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF073780; AAD0186.1; -.
DR INTERPRO: IPR000421; -.
DR INTERPRO: IPR001092; -.
DR PFAM: PF00754; F5_E8_type_C; 1.
DR PROSITE: PS01285; FASBC_1; UNKNOWN_1.
DR SEQUENCE 224 AA; 25545 MW; 45362021000E90E4 CRC64;

Query Match 10.4%; Score 279; DB 11; Length 224;
Best Local Similarity 34.4%; Pred. No. 2.9e-17;
Matches 67; Conservative 32; Mismatches 72; Indels 24; Gaps 7;

QY 298 YPOVCRHRTLMEL---GCEISG-----SEPLGMSGHIDYQITASSIFRTLMD 348
DB 34 YQACKCDQYGANALMSAGATSLDICEPCYHKPLFESGEVTPDQITCS-----NPE 87
QY 349 NF-----TWEPKRLDKQKVNAMTSGHNDOSQWLQVDLVPFTVGTGITQAKDFGHV 403
DB 88 QYGVYSWTANRARNLSQGFCAWLSKYQSSQWLQIDLKEIKYISGLNGHGRDID-- 145
QY 404 QPVSQYKLAISNGE--HMTVYDEKQRKDYFGNFGNDTHRRKNVDPPIYARIRILPP 462
DB 146 EMTVYSYQYTERLWMTIYKQ--TGNNRVFGNSDRSYONLIRPIISRTIRIPL 204
QY 463 SWYGRITLASELGC 477
DB 205 GHWRIATIMELLECC 219

RESULT 11
Q99734 PRELIMINARY; PRT; 263 AA.
AC Q99734;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE NOTCH2 TRANSMEMBRANE PROTEIN (FRAGMENT).
GN NOTCH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN
RN SEQUENCE FROM N.A.
RC TISSUE=T CELL;
RA Lemstra I., Devaux C., Mesnard J.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U77493; AAB19224.1; -.
DR HSSP: P00736; IAPQ.
DR INTERPRO: IPR000083; -.
DR INTERPRO: IPR000152; -.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR000742; -.
DR INTERPRO: IPR001010; -.
DR INTERPRO: IPR001438; -.
DR INTERPRO: IPR001881; -.
DR INTERPRO: IPR001947; -.
DR INTERPRO: IPR002049; -.
DR PFAM: PF00008; EGF_6.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00011; EGFAMININ.
DR PRINTS: PR00012; ENTPEL.
DR PRINTS: PR00286; CHARYD0XIN.
DR PRINTS: PR00287; THIONIN.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; UNKNOWN_7.

```

DR PROSITE; PS01186; EGF_2; 6.
 DR PROSITE; PS01187; EGF_CA; 4.
 KM Glycoprotein; EGF-like domain.
 FT NON TER 1 1
 FT NON TER 263 263
 SO SEQUENCE 263 AA; 28140 MW; 971C7E61C2B7F7B8 CRC64;

Query Match 10.2%; Score 272; DB 4; Length 263;
 Best Local Similarity 26.3%; Pred. No. 1.5e-16;
 Matches 54; Conservative 27; Mismatches 46; Indels 78; Gaps 5;

QY 26 CDPNPCENGICLPGLAVGFSCECPDFTDPNCSSVVEVADSEDEPTAGCPCTPNPCN 85
 DB 63 CSHPCINDTCVDGL--GTYRCSCPLGTGTGKNCQTLVNL-----CSRSPCKN 108
 QY 86 GGT-----CEISEAYRG-----DTFIGYV 104
 DB 109 KGTGVOKKASPOCLCPGAVGACVCPVAVSCDIASRRGLVLEHLQHSVCINAGNTHX 168
 QY 105 CKCRGRNGIHCOHININECEVEPCCKNGICTDLYANISCECPDEPMGRNCQ----- 156
 DB 169 CQCEVGTGTSCEQLDEKASNPQHGATCSDFIGTRCPCVPGYGVNCEVEVDECONQ 228
 QY 157 -----KCSPLGIEG 166
 DB 229 PCQNGGTCTIDLVNHFRCSCPPGTRG 253

RESULT 12

099740 PRELIMINARY; PRT; 263 AA.

AC 099740: 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE SOLUBLE PROTEIN JAGGED (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN.
 RX MEDLINE=95211842; PubMed=7697721;
 RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
 RL "Jagged: a mammalian ligand that activates Notch1.";
 RL Cell 80:909-917(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN.
 RX MEDLINE=97115768; PubMed=8955070;
 RA Limpin A.B., Pepper M.S., McMahon G., Nguyen F., Montesano R.,
 RA Maciag T.;
 RT "An antisense oligonucleotide to the notch ligand jagged enhances
 RT fibroblast growth factor-induced angiogenesis in vitro.";
 RL J. Biol. Chem. 271:32499-32502(1996).
 DR EMBL: U77914; AAC50909.1; -.
 DR HSSP: P00743; ICCF.
 DR INTERPRO: IPR000083; -.
 DR INTERPRO: IPR000152; -.
 DR INTERPRO: IPR000561; -.
 DR INTERPRO: IPR000742; -.
 DR INTERPRO: IPR001881; -.
 DR INTERPRO: IPR002049; -.
 DR PFAM: PF00008; EGF_3.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR PRINTS: PR00012; FNTYPEI.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 2.
 KM Glycoprotein; EGF-like domain.

FT NON TER 1 1
 SO SEQUENCE 263 AA; 28405 MW; E77235CBB360B620 CRC64;

Query Match 10.1%; Score 269.5; DB 4; Length 263;
 Best Local Similarity 36.8%; Pred. No. 2.5e-16;
 Matches 53; Conservative 16; Mismatches 54; Indels 21; Gaps 4;

QY 26 CDPNPCENGICLPGLAVGFSCECPDFTDPNCSSVVEVADSEDEPTAGCPCTPNPCN 85
 DB 6 CLNPCHNGSTCV--VNGSFTCYCKEKGWGPICAQ-----NTNDCSPHCYN 51
 QY 86 GGTCEISEAYRGDTFTGYCKCPGNGIHCOHININECEVEPCCKNGICTDLYANISCEC 145
 DB 52 SGTC-----VDDNM--YRCBCAPGAFGPPCCRIININECSSPCARGATCVDINGRCVC 104
 QY 146 GPEPMGRNCQYKCSPLGIEGII 169
 DB 105 PGHSGAKCOEVSGRPCITMGSVI 128

RESULT 13

055139 PRELIMINARY; PRT; 434 AA.

AC 055139: 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE JAGGED2 PROTEIN (FRAGMENT).
 GN JAG2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Valsecchi V., Ballabio A., Rugari E.I.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y14495; CAA74835.1; -.
 DR HSSP: P00743; ICCF.
 DR INTERPRO: IPR000152; -.
 DR INTERPRO: IPR000561; -.
 DR INTERPRO: IPR000742; -.
 DR INTERPRO: IPR001010; -.
 DR INTERPRO: IPR001438; -.
 DR INTERPRO: IPR001881; -.
 DR INTERPRO: IPR002049; -.
 DR PFAM: PF00008; EGF_9.
 DR PRINTS: PR00010; EGFBLDOD.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR PRINTS: PR00287; THIONIN.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_7.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_10.
 DR PROSITE: PS01186; EGF_2; 7.
 DR PROSITE: PS01187; EGF_CA; 5.
 KM Glycoprotein; EGF-like domain.
 FT NON TER 1 1
 FT NON TER 434 434
 SO SEQUENCE 434 AA; 46176 MW; 09074EB0303205C3 CRC64;

Query Match 10.0%; Score 268.5; DB 11; Length 434;
 Best Local Similarity 35.2%; Pred. No. 6.1e-16;
 Matches 63; Conservative 14; Mismatches 49; Indels 53; Gaps 8;

QY 26 CDPNPCENGIC--LPGLAVGFSCECPDFTDPNCSSVVEVADSEDEPTAGCPCTPNPC 83
 DB 27 CASNPCCANGSCHEVP-----SGFECHCPGSGWSPGC-----ALDIDE-----CASNPC 70
 QY 84 HNGCTCE-----ISEAYRGDT-----FTGYCKCPNGFN 112
 DB 71 AAGGTCTVDYDREGTCICPEQWGWGATCOLDANEGCRKCLNAFCKNLIGYVCDCCLPGWK 130


```

FT      /note="Discoïdin 1/factor VIII-like domain 1"
FT      320..477
FT      /label="Discoïdin-2"
FT      /note="Discoïdin 1/factor VIII-like domain 2"
PN      MO9640769-A1.
PD      19-DEC-1996.
PP      05-JUN-1996; 96MO-US09456.
PR      07-JUN-1995; 95US-0480229.
PA      (PROG-) PROGENITOR INC.
PA      (UYVA-) UNIV VANDERBILT.
PI      Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
DR      WPI: 1997-052233/05.
DR      N-PSDB: T47343.
XX      New developmentally regulated endothelial cell locus-1 (del-1) gene
XX      and conditions involving abnormal angiogenesis
XX      Claim 4; Fig 6; 137pp; English.
XX
XX      Human Del-1 (W10365) is the polypeptide product of the human
XX      developmentally-regulated endothelial cell locus-1 (del-1) gene
XX      (T47343). It shows 94% amino acid homology to the mouse Del-1
XX      protein (W10364). Structurally, members of this novel gene family
XX      contain 1 EGF-like domain and 2 discoïdin 1/factor VIII-like
XX      domains. Del-1 is expressed in endothelial and certain tumour
XX      cells. Its ability to inhibit vascular formation allows its used
XX      as an anti-angiogenic agent. It can be used as a tumour marker,
XX      to identify Del-1 binding partners, and to modulate endothelial
XX      cell growth and blood vessel formation. Recombinant Del-1 can be
XX      produced in transformed host cells utilizing vectors incorporating
XX      del-1 nucleic acids.
XX
XX      Sequence 481 AA:
SQ
Query Match 99.6%; Score 2664; DB 18; Length 481;
Best Local Similarity 99.6%; Pred. No. 9,3e-163;
Matches 479; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 MRSVAVMLVGLSLGVPOFGKGDICDPNFCENGICLPGILAVGSFSCPCDGFDPNCS 60
DB      1 mktsvavvllvgllsivpqlgqdicdpncengiclpqilavgsfscpcdgtfdpncs 60
QY      61 SYVEVASDEEPTSGPCTPNPCNGTCEISEAYRGDTFYGVCPCPRGFNGIHCOHNI 120
DB      61 sveyevsdeepstsgpctpnpcngtceiseayrgdtfilygvcprgfngihcqhni 120
QY      121 NECEVEPCNKGICITDLVANSCECPGEFNGRCQKSGPLGIEGGIISNOOITASSSTH 180
DB      121 necevepcnkgicitdlvanscecpgefngrcqkcsgrplgieggliisngitasssth 180
QY      181 RALFGLOKWPYPYALNKKGILINMTAENDRMKRIQINLORMAVTGVITGAGARIS 240
DB      181 ralfglqkwpypyalnkkgilinawtaendrmrwidnqtrmrvgiltgagrirs 240
QY      241 PEYIKFYKILAYNDKTKWAMKVGYNEDVFRGNIDNNNTYASFPPIKAYVALYQ 300
DB      241 peyikfyklayndktkwamkvgnedvfrgnidnnntyasnfpkayvalyq 300
QY      301 VCCRHCITMELLGELSGCESEPIGMSGHIQDYQITASSIFRLNMDMTWPKARAD 360
DB      301 vccrhcitmellgelsgscsepigmshiqdyqitassifrlndmtwepkrard 360
QY      361 KQGVNAWTSQHNDSQWLVQDLVLPKVTGIIIGCANDFGHVQFVSYKLAVSNDGEHW 420
DB      361 kqgvnawtsghndsqwlvqdlvlpkvtgiiigcandfghvofvsyklavsndgehw 420

```

```

DB      361 kqgvnawtsghndsqwlvqdlvlpkvtgiiigcandfghvfygsklaynsndgehw 420
QY      421 TVYQDQKQKXKXVPOGNDNDTFRKXVNDPPIYARHITILPMWYGGITLASELCTEE 480
DB      421 tvyqdkqkxkvfgnfdndtfrkxvndppiyarhltlilpmwyggitlaseelctee 480
QY      481 E 481
DB      481 e 481
DB      481 e 481
XX
XX      RESULT 2
XX      ID W10364 standard; Protein; 480 AA.
XX      W10364;
XX      03-MAY-1997 (first entry)
XX      Mouse developmentally-regulated endothelial cell locus-1 protein.
XX      Del-1; developmentally-regulated endothelial cell locus-1;
XX      signal transduction; cancer; tumour marker; angiogenesis;
XX      diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX      cardiac ischaemia; stroke; vascular disease; wound healing;
XX      vulnerability; bone formation; diagnosis; therapy.
XX
XX      Mus sp.
XX
XX      Key location/Qualifiers
XX      Peptide 1..23
XX      Protein /label="Sig-peptide
XX      24..480
XX      Domain /label="Mat-protein
XX      26..59
XX      /label="EGF-1
XX      /note="Epidermal growth factor-like domain 1"
XX      78..116
XX      /label="EGF-2
XX      /note="Epidermal growth factor-like domain 2"
XX      123..154
XX      /label="EGF-3
XX      /note="Epidermal growth factor-like domain 3"
XX      158..314
XX      /label="Discoïdin-1
XX      /note="Discoïdin 1/factor VIII-like domain 1"
XX      319..476
XX      /label="Discoïdin-2
XX      /note="Discoïdin 1/factor VIII-like domain 2"
XX
XX      MO9640769-A1.
XX      19-DEC-1996.
XX      05-JUN-1996; 96MO-US09456.
XX      07-JUN-1995; 95US-0480229.
XX      (PROG-) PROGENITOR INC.
XX      (UYVA-) UNIV VANDERBILT.
XX      Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
XX      WPI: 1997-052233/05.
XX      N-PSDB: T47338.
XX      New developmentally regulated endothelial cell locus-1 (del-1) gene
XX      - used to develop prods. for the diagnosis and treatment of cancer
XX      and conditions involving abnormal angiogenesis
XX      Claim 3; Fig 6; 137pp; English.
XX
XX      Murine Del-1 (W10364) is the polypeptide product of the murine

```



```

Db      420 mwbgqekqrkdkvfgsnfndthrkknvldppiyarfirllpwsygrllrsellgcaee 479
Qy      481 E 481
        480 e 480

RESULT  4
W94697  standard; Protein: 321 AA.
ID      W94697
AC      W94697;
XX      05-MAY-1999 (first entry)
DE      Human milk fat globule protein MFG-E8.
XX      Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX      discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
XX      diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX      angiogenesis.
XX      Homo sapiens.
XX      US587281-A.
XX      02-MAR-1999.
XX      05-JUN-1996; 96US-0659235.
XX      05-JUN-1996; 96US-0659235.
XX      07-JUN-1995; 95US-0480229.
XX      (PROG-) PROGENITOR INC.
XX      (UYVA-) UNIV VANDERBILT.
XX      Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
XX      WPI; 1999-189720/16.
XX      Del-1 polypeptide sequences - useful for treatment of cancer,
XX      diabetic retinopathy, rheumatoid arthritis and endometriosis
XX      Example: Column 63-64; 73pp; English.
XX      The present sequence represents human milk fat globule protein MFG-E8,
XX      which has homology to the developmentally-regulated endothelial cell
XX      locus 1 (del-1). The Del-1 protein has epidermal growth factor (EGF) like
XX      domains and discoidin I/factor VIII-like domains. The Del-1 proteins
XX      have an inhibitory effect on angiogenesis (blood vessel growth). This
XX      activity may be useful clinically to prevent neovascularisation of
XX      tissues such as tumour nodules and prevention of metastases. The anti-
XX      angiogenic activity of Del-1 may be used to treat abnormal conditions
XX      that result from angiogenesis, including cancer, diabetic retinopathy,
XX      rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX      angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX      stroke, wound healing and peripheral vascular disease. Del-1 is also
XX      useful for promoting bone formation. Del-1 binds to alpha V beta 3
XX      integrin, and is an apoptosis inducer.
XX      Sequence 321 AA:

Query Match      64.2%; Score 1717; DB 20; Length 321;
Best Local Similarity 98.4%; Pred. No. 1,8e-102;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Db      61 qinlqkrmrvtgyltgakrlispeyikfyklayndgkltwamykvkglmednvtfrgnid 120
Qy      278 NNTPEANSFTPIKAQYVRLYPOVCRRHCTLRMELLGCELSGCSPELGMKSGHIDYOIT 337
        121 nntpyansftppikagylvrlpvcrrhctlrme1lgcelsgcsepiymksg1ldyqlt 180
Qy      338 ASSIFRTIMNDMEFTWEPRKARLDKOGKVNAMTSGHNDQSOMLQYDLVYPRYKVGIIITOGA 397
        181 assifrltmdmftwepkrarldkqgkvnawtsgndgsqvlvxlivpkyvgi1ltgga 240
Qy      398 KDFGHVQFVSGYKLAYSNDGEHMTVYODEKQRDKVPOGNFNDNTHRKKNVDPPIYARHI 457
        241 kdfghvqfvgysklaysndgehwetvxdgkqrkdkvsgnfdndthrkknvldppiyarhi 300
Db      458 RILPWSYGRITLASLGLCT 478
        301 rilpwsygritlasellgct 321

RESULT  5
W10366  standard; Protein: 221 AA.
ID      W10366
AC      W10366;
XX      03-MAY-1997 (first entry)
DE      Murine Del-1 truncated minor protein.
XX      Del-1; developmentally-regulated endothelial cell locus-1;
XX      signal transduction; cancer; tumour marker; angiogenesis;
XX      diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX      cardiac ischaemia; stroke; vascular disease; wound healing;
XX      vulnerability; bone formation; diagnosis; therapy.
XX      Mus sp.
XX      Location/Qualifiers
XX      key 1..23
XX      FT /label= Sig-peptide
XX      FT Protein 24..480
XX      FT /label= Mat-protein
XX      FT Domain 26..59
XX      FT /label= EGF-1
XX      FT /note= "epidermal growth factor-like domain 1"
XX      FT Domain 78..116
XX      FT /label= EGF-2
XX      FT /note= "epidermal growth factor-like domain 2"
XX      FT Domain 123..154
XX      FT /label= EGF-3
XX      FT /note= "epidermal growth factor-like domain 3"
XX      FT Domain 158..221
XX      FT /label= Discoidin-1
XX      FT /note= "truncated discoidin I/factor VIII-like domain 1"
XX      W09640769-A1.
XX      19-DEC-1996.
XX      05-JUN-1996; 96WO-US09456.
XX      07-JUN-1995; 95US-0480229.
XX      (PROG-) PROGENITOR INC.
XX      (UYVA-) UNIV VANDERBILT.
XX      Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
XX      WPI; 1997-052233/05;
XX      N-PSDB; 747339.
XX      New developmentally regulated endothelial cell locus-1 (del-1) gene

```

PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis

XX Example: Fig 12, 137pp; English.

XX A truncated version (W10366) of murine developmentally-regulated
 CC endothelial cell locus-1 (del-1) gene product (see also W10364)
 CC contains a signal peptide, all 3 EGF-like domains but only a
 CC partial N-terminal discoidin I/factor VIII-like domain (about 40%).
 CC It is the product of a murine del-1 minor sequence (747339). This
 CC transcript was cloned only from mouse embryonic libraries, but was
 CC verified through cloning of several independent cDNAs.

XX Sequence 221 AA:

Query Match 42.6%; Score 1139.5; DB 18; Length 221;
 Best Local Similarity 90.5%; Pred. No. 9e-66;
 Matches 200; Conservative 6; Mismatches 14; Indels 1; Gaps 1;

QY 1 MRSVAVMLVGLSLGVPOFGKGDICDPNCENGICLPGLAVGSFSCPCDPTDNCPS 60
 DB 1 mhlvaawllvylslgypgfygkylcpncpceagylclsgladdsfsccepfagpncs 60
 QY 61 SYVEVASDEEPTAGCPNPNCHNGCTEISEAYRGDPFIQVCKCPRGNGIHCQHN 120
 DB 61 syvevasdeektsagcpnncpncngctcseayrgdftfygvcckpgrngihcqhnl 120
 QY 121 NCEVEPCNKGICITDLYANVSCPCGPFMGRNCOYCSGPIGIEGIIISNOITASTH 180
 DB 121 ncevepcnrgictldlyanvscpcgpfmrgncqscghlqiegllsnqitassnh 180
 QY 181 RALPFLQKWPYRRLNKKGLINAMTAENRMRKRWIOINL 221
 DB 181 ralflgkwprrlknkglinamtaendmrp-wlqvtrv 220

RESULT 6
 Y94454
 ID Y94454 standard; Protein: 426 AA.

XX AC Y94454;
 XX 11-SBP-2000 (first entry)
 XX Mouse lactadherin protein.
 XX Human; lactadherin; MGF-E8; anti-tumour; immune response;
 XX exosome; dendritic cell.
 XX OS Mus sp.
 XX Key Location/Qualifiers
 XX Peptide 1..22
 XX Protein 23..426
 XX /label= Secretion_signal
 XX /label= Lactadherin
 XX Binding-site 87..89
 XX /label= Integrin_binding_site
 XX Misc-difference 93..111
 XX /note= "encoded by GT"
 XX
 XX EP1004664-A1.
 XX
 XX 31-MAY-2000.
 XX
 XX 24-NOV-1998; 98EP-0402925.
 XX
 XX 24-NOV-1998; 98EP-0402925.
 XX
 XX (INRM) INSEEM INST NAT SANTE & RECH MEDICALE.
 XX (CURT-) INST CURIE.
 XX

DR WPI: 2000-352597/31.
 DR N-FSDB: A27141.

XX Chimeric isolated (human) lactadherin polypeptide that functions as an
 PT adaptor of cross-priming to eliminate pathogenic antigens; e.g. in
 PT tumours -
 XX

XX Disclosure; Page 12; 20pp; English.

XX Lactadherin protein was found in exosomes produced by dendritic cells.
 CC The protein is involved in the phagocytosis of particulate antigens by
 CC dendritic cells. Exosomes produced by dendritic cells exposed to
 CC tumour antigens induce potent immune responses. Lactadherin or variants
 CC of it may be used in the mediation of an immune response. Variants of
 CC lactadherin may be used for inhibition and/or stimulation of the
 CC cross-priming of antigens and stimulation of the phagocytosis of
 CC antigens by dendritic cells. Compositions derived from lactadherin can
 CC also be used to monitor an immune response, more specifically a CTL
 CC (cytotoxic T-lymphocyte) response and also to produce CTLs specific
 CC for a selected antigen. The present sequence is the mouse
 CC lactadherin protein.

XX Sequence 426 AA:

Query Match 42.4%; Score 1134.5; DB 21; Length 426;
 Best Local Similarity 47.4%; Pred. No. 3.7e-65;
 Matches 216; Conservative 67; Mismatches 118; Indels 55; Gaps 7;

QY 23 GDIDPNPCNGGICLPGLAVGSFSCPCDPTDNCSSVVEVASDEEPTAGPCTPNP 82
 DB 23 gdfdsalclngstclty-qndnlyclcpqeflglvcne-----tergpcpnp 72
 QY 83 CHNGCTEIS-EAYRGDPFIQVCKCPRGNGIHCQHNINECEVPCKNGICTDLYANV 141
 DB 73 cyndakclvldtqrgdltleylcpqevygsjh----- 106
 QY 142 SCECPGFMGRNCOYCSGPIGIEGIIISNOITASTHRAFLGLOKWPYRRLNKKGL 201
 DB 107 -----celctcstqlgmegqaladsglsasyymgmfmqtwgdelarjyrlgl 154
 QY 202 INAMTAENRMRKRWIOINLQKRWYGVITTOGAKRIGSPYIKFKIAYSNCKRWAMY 261
 DB 155 vnaahnasnyds-kpwrlqvnllrkmsvsgmtqgsaagragaejklrkvayslgdrifeti 213
 QY 262 KVKGTNEDMYFRGNIDNNTPIANSFTPIKAQYRRLPYOVCRHCTLMELICEISGCS 321
 DB 214 qde-sggdkeftlgndnslkvmtnpdleagylrlpvschrgctrltelgcelngcl 272
 QY 322 EPLCMKSGHIQDYQITASIFERTLMDFTWEPFRKARLDKQKYNAMTSGHNOQWLOV 381
 DB 273 eplglknltlpdsqmasassyltwnlrafgwphylgrldngqklawlaqsnakawldv 332
 QY 382 DLAVPTKVTGIIIOGAKDEGAVQEVGSYKLAYSNDSEHWTVYODEKORCKDVFGNFND 441
 DB 333 dlqctqvtglitlgardghlqyvesykvahsdvqvtvy--eeqssavltqgndnn 390
 QY 442 THRNKVIDPEYVARHRIIPMSWYGRITLASDELIC 477
 DB 391 shknfkexkfmariyrvlpvswmrtilrlteljgc 426

RESULT 7
 W94684
 ID W94684 standard; Protein: 221 AA.
 XX
 XX W94684;
 XX
 XX 05-MAY-1999 (first entry)
 XX
 XX Truncated murine Del-1 protein.
 XX
 XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 XX

CC XX OS Mus sp.
CC XX PN US5877281-A.
CC XX PD 02-MAR-1999.
CC XX PE 05-JUN-1996; 96US-0659235.
CC XX PR 05-JUN-1996; 96US-0659235.
CC XX PR 07-JUN-1995; 95US-0480229.
CC XX PA (PROC-) PROGENITOR INC.
CC XX PA (UYVA-) UNIT VANDERBILT.
CC PI Hogan B., Quartermous T., Snodgrass HR., Zupancic TJ;
CC DR WPI: 1999-189720/16.
CC DR N-PSTDB: X18507.
CC PT Del-1 polypeptide sequences - useful for treatment of cancer,
CC PT diabetic retinopathy, rheumatoid arthritis and endometriosis
CC PS Claim 3; Column 71-72; 73pp; English.

CC XX The present sequence is truncated murine developmentally-regulated
CC endothelial cell locus 1 (Del-1). The protein has epidermal growth factor
CC (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1
CC proteins have an inhibitory effect on angiogenesis (blood vessel growth),
CC this activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC Integrin, and is an apoptosis inducer.

CC XX Sequence 221 AA:

Query Match 42.4%; Score 1133.5; DB 20; Length 221;
Best Local Similarity 90.0%; Pred. No. 2,2e-65;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

OY 1 MKRSVAVMLVLGLSLGPOGKGDICDPNPENGSGICDLPLGVASFSECPDGPFPNCS 60
Db 1 mkhlvaavallvglslygpdqfgydclcpnpengsgicldgladstfsceepgfapnps 60

OY 61 SVYEVAASDEEPTAGPCPTPNCCHNGSGICELSEAYRSDPFITGYCKCRPGNIGHOQMI 120
Db 61 svyevaaadeeptasagpc.lpnchngmgfcelseayrgdfitfyvckcprfgnhqomhi 120

OY 121 NEEVEPEPKNGICTDI.VANYSCGPCRGEMRNQYKCGSPGIRGSITSNOQTASSRH 180
Db 121 neeeveperngictdi.vanyscepcrgsfingrcnykcsghilsggi.lsnqqtassrh 180

OY 181 RALFLQLQWRYPYRARLKKGIINAMTAAENRKMKIQLNL 221
Db 181 ralflqlqwrypyralkrkylinaatlaendrrpw-wlqvltv 220

RESULT 8
R77252
ID R77252 standard; Protein: 387 AA.
XX R77252:
XX 21-NOV-1995: (first entry)
XX JT

```

DE HMFG 46 kDa antigen.
XX
XX
KW HMFG; human milk fat globule; antigen; immunogen; vaccine; virucide;
KV epithelium; tumor; breast cancer; monoclonal antibody; Mab.
XX
XX OS Homo sapiens.
PM WC95J5171-A.
PD 08-JUN-1995.
XX
XX PF 05-DEC-1994; 94MO-US13967.
XX
XX PR 03-DEC-1993; 93US-0162402.
XX
XX PA (CANC-) CANCER RES FUND CONTRA COSTA.
XX
XX PI Ceriani RL, Larocca DJ, Peterson JA;
DR WPI: 1995-215151/28.
DR N-PDB: Q91198.
XX
XX PT 46 kD apparent molecular weight human milk fat globule antigen -
PT used in assays to determine the presence of a carcinosarcoma of
XX epithelial origin, and in a vaccine against neoplastic tumors
XX
XX PS Claim 6: Page 46-47, 68pp: English.
CC
CC A complete cDNA sequence for the 46 kDa HMFG antigen, a major
CC component of the apical surface of the normal breast epithelial
CC cell, was obtained by PCR and RACE methods. cDNA clones can
CC be used to prepare MAbs for use in immunotherapy. Immunohistochemistry,
CC prognosis, diagnosis, imaging and therapy. Recombinant antigen can
CC be expressed in prokaryotic or (glycosylated) in eukaryotic cells.
XX
XX Sequence 387 AA:
SQ
Query Match 38.1%; Score 1020.5; DB 16; Length 387;
Best Local Similarity 48.8%; Pred. No. 64e-58;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5
OY 78 CTGTCGGCGTGC-EISAVRGDFEIGVYCKCPRGNGHCOHNINCEVEPEKKGICTD 136
OY I : |||||I::|||: ||| : | : : : :
Db 27 CSMNPHOGGLCOELISEGVGVDFIPSYCTCLKYAGNH----- 65
OY 137 LVANYSCEEPGFEMGRNCQYKCSGPSTGICGTISNOOTATSTNRALFGLOKMYUAYRL 196
OY -----CECKCVETPRGMENGLANSQSAAVSIVFLIDHWPELARI 108
OY 197 NRKKLTNAWTAENDRKRKMTIOINLOKMKRYGVITGSARKISPEYIKFYLIANSNGK 256
OY I : |||||I:: || |||: ||| |||||I:: : ||| : : ||| : ?
Db 109 NRGWMAEWLPESNDP-NPWLYGNLLRTMWVGVTGASIASHEYLAFKVEYSLGH 167
OY 257 TW-AMKKYSGNEEDWFRGNINDNTGYASFPRIKDYARLYPOUGSRNCTIRMEILGC 315
OY I : |||||I:: ||| : ||| : : : : : |||||I:: ||| :
Db 168 EFFLDLVKKKHKEIV-GMKNKAVALHLETPREAGVYEVLRPSCHACILFEILLGC 225
OY 316 ELISGCEPFLGMSGHIODYDTASSIFRLAMDHTWEPRKALDIQGVNANMTSNNHQ 375
OY I : |||||I:: ||| : ||||| : : : : |||||I:: ||| :
Db 226 EINGCENPGLIMSPDKQTASSYKCKWGTHLSWPSADLIDGQNLAWAGSYGN 285
OY 376 SQWLQVDLVLPKTKYGIITGAKDEFCHVOVFQSTRKLAYSNDGEHWTYVODEKORKDVQ 435
OY 286 DQQLYGVDAISSKEYGVGLTGGAINTSVTFVAASYVUYVASDNANSELEYDPRTGSKILFP 345
OY 436 GNFDNDTRKKNVIDPPYARHINILFWMSYGRITLASBLIGC 477
OY I : |||||I:: : |||||I:: : |||||I:: : |||||I::
Db 346 GWGDASHSHKNIIEPLIARYVLIIPIVWAHRIALRIELLEIGC 477

```

ID	Accession	Standard	Protein	Length
Y94453	Y94453	standard	Protein: 387 AA.	
XX	Y94453			
AC	Y94453			
DT	11-SEP-2000	(first entry)		
DE	Human lactadherin protein.			
XX				
XX	Human: Lactadherin; MGR-BB; anti-tumour; immune response;			
KM	exosome; dendritic cell.			
XX				
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
FT	Peptide	1..23		
FT	Protein	/label= Secretion_signal		
FT	Binding-site	24..387		
FT		/label= Lactadherin		
FT		46..48		
FT		/label= Integrin_binding_site		
XX				
PN	EP1004664-A1.			
PD	31-MAY-2000.			
XX				
PE	24-NOV-1998;	98EP-0402925.		
PR	24-NOV-1998;	98EP-0402925.		
XX				
PA	(INNM) INSERM INST NAT SANTE & RECH MEDICALE.			
PA	(CURT-) INST CURIE.			
XX				
DR	WPI: 2000-352597/31.			
DR	N-PSDB: A27140.			
XX				
XX	Chimeric isolated (human) lactadherin polypeptide that functions as an			
XX	adaptor of cross-priming to eliminate pathogenic antigens, e.g. in			
XX	tumors -			
XX				
XX	Example 3; Page 12; 20pp; English.			
XX				
CC	Lactadherin protein was found in exosomes produced by dendritic cells			
CC	The protein is involved in the phagocytosis of particulate antigens by			
CC	dendritic cells. Exosomes produced by dendritic cells exposed to			
CC	tumour antigens induce potent immune responses. Lactadherin or variant			
CC	of it may be used in the mediation of an immune response. Variants of			
CC	Lactadherin may be used for inhibition and/or stimulation of the			
CC	cross-priming of antigens and stimulation derived from the phagocytosis of			
CC	antigens by dendritic cells. Compositions derived from lactadherin can			
CC	also be used to monitor an immune response, more specifically a CTL			
CC	(cytotoxic T-lymphocyte) response and also to produce CTLs specific			
CC	for a selected antigen. The present sequence is the human			
CC	lactadherin protein.			
XX				
SO	Sequence 387 AA:			
Query Match	38.1%; Score 1020.5; DB 21; Length 387;			
Best Local Similarity	48.8%; Pred. No. 6.4e-58;			
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps				
Q7	78 CIPNNGHGNGC-EISEADRDPTIGYCKCRGPNNGIHCNHNINCEVECKNGGICLD 136			
DB	27 csknpchnglceiseigevrgdvfpystctclkygnh----- 65			
Q7	137 LVANTSCSCPEPEWGNQCYKCSGPLIEBGLISMOQTASSTHRRLRGLQKWPYIARL 196			
DB	66 -----cckcveplisemngnlanslaasvrvflglqlmwpealari 108			
Q7	197 NKKGLINATWADENRKMFTQIQLQKKRVTGVITGCAKRICSPYIKETKIASMDQG 256			
DB	109 nrgamwawpssndd-npwldvnlitrmwvlgvvtgqasrlasheykativayslngn 167			

[illegible]

XX Example; Page 71-72; 137pp; English.
 PS
 CC A discoidin-like domain (W10367) was identified in mouse
 CC developmentally-regulated endothelial cell locus-1 (Del-1)
 CC sequences (see also W10364) derived from a trapped exon and
 CC mouse embryo cDNAs. It showed homology to discoidin-like domains
 CC from human milk fat globule protein, human factor V, mouse factor
 CC VIII, X-A5b1 and X-A5b2 (domains of xenopus neuronal antigen A5)
 CC and discoidin I.
 XX
 SQ Sequence 85 AA;
 Query Match 14.5%; Score 387; DB 18; Length 85;
 Best Local Similarity 91.5%; Pred. No. 3.7e-18;
 Matches 75; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
 QY 382 DLVPTKVTGTTG--AKDEGHVGFVGSYKLAISNDEHWTYYODEKQKRVQGNFD 439
 Db 1 dlvpkvtgltcgxxakdrgvltvgsyxlaysndghmwmvhdqkqrkdvfgnfd 60
 QY 440 NDTHRKNVIDPPYIARHRIPL 461
 Db 61 ndthrkvnldppiyarfrilp 82
 RESULT 15
 W94689
 ID W94689 standard; peptide: 85 AA.
 XX
 AC W94689;
 XX
 DT 05-MAY-1999 (first entry)
 XX
 DE Discoidin I/factor VIII like domain #1.
 XX
 KM Del-1: developmentally-regulated endothelial cell locus 1; cancer;
 KM discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
 KM diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KM angiogenesis.
 XX
 OS Mus sp.
 OS Synthetic.
 XX
 PN US5877281-A.
 XX
 PD 02-MAR-1999.
 XX
 PF 05-JUN-1996; 960S-0659235.
 XX
 PR 05-JUN-1996; 960S-0659235.
 PR 07-JUN-1995; 950S-0480229.
 XX
 XX (PROG-) PROCENTOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
 XX
 DR WPI; 1999-189720/16.
 XX
 PT Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Example; Column 37-38; 73pp; English.
 XX
 CC The present sequence represents a discoidin I/factor VIII-like domain.
 CC The present invention describes developmentally-regulated endothelial
 CC cell locus 1 (Del-1). Del-1 has epidermal growth factor like domains and
 CC discoidin I/factor VIII-like domains. The Del-1 proteins have an
 CC inhibitory effect on angiogenesis (blood vessel growth); this activity
 CC may be useful clinically to prevent neovascularisation of tissues such
 CC as tumour nodules and prevention of metastases. The anti-angiogenic

CC activity of Del-1 may be used to treat abnormal conditions that result
 CC from angiogenesis, including cancer, diabetic retinopathy, rheumatoid
 CC arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be
 CC used to treat cardiac ischaemia, thrombotic stroke, wound healing and
 CC peripheral vascular disease. Del-1 is also useful for promoting bone
 CC formation. Del-1 binds to alpha v beta 3 integrin, and is an apoptosis
 CC inducer.
 XX
 SQ Sequence 85 AA;
 Query Match 14.5%; Score 387; DB 20; Length 85;
 Best Local Similarity 91.5%; Pred. No. 3.7e-18;
 Matches 75; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
 QY 382 DLVPTKVTGTTG--AKDEGHVGFVGSYKLAISNDEHWTYYODEKQKRVQGNFD 439
 Db 1 dlvpkvtgltcgxxakdrgvltvgsyxlaysndghmwmvhdqkqrkdvfgnfd 60
 QY 440 NDTHRKNVIDPPYIARHRIPL 461
 Db 61 ndthrkvnldppiyarfrilp 82

Search completed: May 23, 2001, 08:38:08
 Job time: 129 sec

Wed May 23 08:50:26 2001

us-09-237-981-14_copy_33_513.clo.ram

:age 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:37:29 ; Search time 48.45 Seconds

(without alignments)
1596.925 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513

Perfect score: 2675

Sequence: 1 MKRSVAVMLVGLSLGVDPF.....MSWYGRITLSELLCTEE 481

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 932693

Minimum DB seq length: 0

Maximum DB seq length: 481

Post-processing: Minimum Match 08

Maximum Match 1008

Database :

Listing first 45 summaries

Pending Patents: AA.Maln.*
1: /cgn2_6/ptodata/2/paa/US05.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2646	98.9	481	10	US-08-659-235B-29
2	2519.5	94.2	480	8	US-08-480-229B-10
3	2519.5	94.2	480	10	US-08-659-235B-10
4	2519.5	94.2	480	16	US-09-237-981-10
5	1717	64.2	321	8	US-08-480-229B-21
6	1717	64.2	321	10	US-08-659-235B-21
7	1717	64.2	321	16	US-09-237-981-21
8	1406.5	52.6	23	23	US-60-207-315-526
9	1406.5	52.6	23	23	US-60-207-315-526
10	1405.5	52.5	311	23	US-60-230-435-1519

11	1156	43.2	463	19	US-09-582-340-4	Sequence 4, Appl
12	1133.5	42.9	221	16	US-09-237-981-29	Sequence 29, Appl
13	1039.5	38.9	203	23	US-60-233-644-77	Sequence 17, Appl
14	1036.5	38.7	185	23	US-60-230-435-1806	Sequence 1606, Ap
15	1020.5	38.1	387	19	US-09-582-340-2	Sequence 2, Appl
16	948	35.4	320	8	US-08-480-229B-20	Sequence 20, Appl
17	948	35.4	320	10	US-08-659-235B-20	Sequence 20, Appl
18	948	35.4	320	16	US-09-237-981-20	Sequence 20, Appl
19	643	24.0	217	8	US-08-482-596A-3	Sequence 3, Appl
20	643	24.0	217	8	US-08-482-596A-2	Sequence 4, Appl
21	643	24.0	218	8	US-08-482-596A-1	Sequence 1, Appl
22	475.5	17.8	218	8	US-08-480-229B-1	Sequence 1, Appl
23	387	14.5	83	8	US-08-959-222B-1	Sequence 1, Appl
24	387	14.5	83	10	US-09-237-981-11	Sequence 1, Appl
25	387	14.5	83	16	PCT-US00-22226A-11	Sequence 1, Appl
26	373	13.9	159	1	PCT-US00-22226A-11	Sequence 11, Appl
27	373	13.9	159	1	PCT-US00-22226A-11	Sequence 25, Appl
28	368	13.8	109	8	US08-463-2226A-2	Sequence 8, Appl
29	364.5	13.6	160	1	PCT-US00-22226A-8	Sequence 165, App
30	364.5	13.6	160	1	PCT-US00-22226A-8	Sequence 165, App
31	351	13.1	161	23	US-60-208-043-215	Sequence 9, Appl
32	351	13.1	161	17	US-09-331-793-9	Sequence 9, Appl
33	344.5	12.8	160	1	PCT-US00-22226A-9	Sequence 9, Appl
34	343.5	12.8	160	1	PCT-US00-22226A-9	Sequence 9, Appl
35	332.5	12.5	160	1	PCT-US00-22226A-10	Sequence 10, Appl
36	332.5	12.5	160	1	PCT-US00-22226A-10	Sequence 12, Appl
37	332.5	12.5	150	1	PCT-US00-22226A-12	Sequence 12, Appl
38	323	12.1	150	1	PCT-US00-22226A-12	Sequence 12, Appl
39	323	12.1	150	1	PCT-US00-22226A-12	Sequence 3075, Ap
40	319.5	11.9	62	23	US-60-192-739-3075	Sequence 165, App
41	317	11.9	126	23	US-60-233-940-165	Sequence 165, App
42	307	11.5	147	23	US-60-185-361-589	Sequence 589, App
43	304	11.4	110	8	US-08-463-294-24	Sequence 24, Appl
44	301.5	11.3	96	23	US-60-169-840-5111	Sequence 5111, Ap
45	301.5	11.3	96	23	US-60-169-867-4347	Sequence 4347, Ap

ALIGNMENTS

RESULT 1
US-08-659-235B-29
: Sequence 29, Application US/08659235B
: GENERAL INFORMATION:
: APPLICANT: Overtemous, Thomas
: APPLICANT: Hogan, Bridg
: APPLICANT: Snodgrass, H. Ralph
: APPLICANT: Zupancic, Thomas J.
: TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/659,235B
: FILING DATE: 05-JUN-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Polissant, Brian M.
: REGISTRATION NUMBER: 6462
: REFERENCE/DOCKET NUMBER: 8907-034
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090

Wed May 23 08:50:26 2001

us-09-237-981-14_copy_33_513.clo.ram

Page 2

TELEFAX: (212) 869-8864/9741
TELEX: 66141 Penile
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-659-235B-29

Query Match
Best Local Similarity 98.98; Score 2646; DB 10; Length 481;
Matches 477; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNFCENGICLPGLAVGSFSCCEPDGFTDPNCS 60
DB 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNFCENGICLPGLAVGSFSCCEPDGFTDPNCS 60
QY 61 SYVEVASDEEPTASGCTPRPCHNGCTEISEAYRGDPFTGYVCKCPRGNGIHCOHNI 120
DB 61 SYVEVASDEEPTASGCTPRPCHNGCTEISEAYRGDPFTGYVCKCPRGNGIHCOHNI 120
QY 121 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPIGEGGIIISNOITASSH 180
DB 121 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPIGEGGIIISNOITASSH 180
QY 181 RALFGLOKWPYARLANKKGLINAMTAENDRMKRMQIOLQKRMVGVITOGAKRIGS 240
DB 181 RALFGLOKWPYARLANKKGLINAMTAENDRMKRMQIOLQKRMVGVITOGAKRIGS 240
QY 241 PEYIKFYKIAVNSDKGTWAMYKVGKNEEDMVEFRGNIDNTPYANSFTPIKAQYVRLYPQ 300
DB 241 PEYIKFYKIAVNSDKGTWAMYKVGKNEEDMVEFRGNIDNTPYANSFTPIKAQYVRLYPQ 300
QY 301 VCRHCTRLMELLCGELSGCSEPLGKSGHIDYQITASSIFRTLMQDMFTWEPKARLD 360
DB 301 VCRHCTRLMELLCGELSGCSEPLGKSGHIDYQITASSIFRTLMQDMFTWEPKARLD 360
QY 361 KQGVNANTSGHNDOSQWLQVLDLVPKVTGIIITOGAKDFGHVQFVGSYKLAYSNDGEHW 420
DB 361 KQGVNANTSGHNDOSQWLQVLDLVPKVTGIIITOGAKDFGHVQFVGSYKLAYSNDGEHW 420
QY 421 TVYODEKORKDKVPOGFNDNTHRKKNVIDPPIYARHRIILPMSWGRITLASSELLCTEE 480
DB 421 TVYODEKORKDKVPOGFNDNTHRKKNVIDPPIYARHRIILPMSWGRITLASSELLCTEE 480
QY 481 E 481
DB 481 E 481

RESULT 2
US-08-480-229B-10
Sequence 10, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quartermours, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229B-10

Query Match
Best Local Similarity 94.28; Score 2519.5; DB 8; Length 480;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;
QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNFCENGICLPGLAVGSFSCCEPDGFTDPNCS 60
DB 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNFCENGICLPGLAVGSFSCCEPDGFTDPNCS 60
QY 61 SYVEVASDEEPTASGCTPRPCHNGCTEISEAYRGDPFTGYVCKCPRGNGIHCOHNI 120
DB 61 SYVEVASDEEPTASGCTPRPCHNGCTEISEAYRGDPFTGYVCKCPRGNGIHCOHNI 120
QY 121 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPIGEGGIIISNOITASSH 180
DB 121 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPIGEGGIIISNOITASSH 180
QY 181 RALFGLOKWPYARLANKKGLINAMTAENDRMKRMQIOLQKRMVGVITOGAKRIGS 240
DB 181 RALFGLOKWPYARLANKKGLINAMTAENDRMKRMQIOLQKRMVGVITOGAKRIGS 240
QY 241 PEYIKFYKIAVNSDKGTWAMYKVGKNEEDMVEFRGNIDNTPYANSFTPIKAQYVRLYPQ 300
DB 241 PEYIKFYKIAVNSDKGTWAMYKVGKNEEDMVEFRGNIDNTPYANSFTPIKAQYVRLYPQ 300
QY 301 VCRHCTRLMELLCGELSGCSEPLGKSGHIDYQITASSIFRTLMQDMFTWEPKARLD 360
DB 301 VCRHCTRLMELLCGELSGCSEPLGKSGHIDYQITASSIFRTLMQDMFTWEPKARLD 360
QY 361 KQGVNANTSGHNDOSQWLQVLDLVPKVTGIIITOGAKDFGHVQFVGSYKLAYSNDGEHW 420
DB 361 KQGVNANTSGHNDOSQWLQVLDLVPKVTGIIITOGAKDFGHVQFVGSYKLAYSNDGEHW 420
QY 421 TVYODEKORKDKVPOGFNDNTHRKKNVIDPPIYARHRIILPMSWGRITLASSELLCTEE 480
DB 421 TVYODEKORKDKVPOGFNDNTHRKKNVIDPPIYARHRIILPMSWGRITLASSELLCTEE 480
QY 481 E 481
DB 481 E 480

RESULT 3
US-08-659-235B-10
Sequence 10, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quartermours, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

```

TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-10

```

```

Query Match          94.2%; Score 2519.5; DB 10; Length 480;
Best Local Similarity 94.2%; Pred. No. 3.5e-208;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCENGGICLPLGLAVSFCSECPDGFTPNCS 60
DB 1 MKHLVAMLLVGLSLGVPOFGKGDICPNPCENGGLSLGLADDSFSCDEPGEFAPNCS 60
QY 61 SVEVVASDEEPTSGAGPCTPMPCHNGSTCEISEAYRGDTFGLGYCKCPRGFGNCHQHN 120
DB 61 SVEVVASDEEPTSGAGPCTPMPCHNGSTCEISEAYRGDTFGLGYCKCPRGFGNCHQHN 120
QY 121 NECEVEPCKNGGICIDLVANYSCECPGEFGRNCOYKCSGGLIGGIIISNOGITASSSTH 180
DB 121 NECEVEPCKNGGICIDLVANYSCECPGEFGRNCOYKCSGGLIGGIIISNOGITASSSTH 180
QY 181 RALFGLQKWPYARLANKKGLINAMTAENDRKRWIOINLQRMKRYGVITOGAKRIGS 240
DB 181 RALFGLQKWPYARLANKKGLINAMTAENDRWP-WIOINLQRMKRYGVITOGAKRIGS 239
QY 241 PEYIKFKTIAYSNDGKTWAMTKVAGTNEDEVFRGNIDNNTPYANSFPPPIKAQYVALYQ 300
DB 240 PEYIKFKTIAYSNDGKTWAMTKVAGTNEDEVFRGNIDNNTPYANSFPPPIKAQYVALYQ 299
QY 301 VCRHCHLRLMELGELSGCSEPIGMKSGHIOIYOTASSITPRLTNDMTWEPKRRARD 360
DB 300 ICRHCHLRLMELGELSGCSEPIGMKSGHIOIYOTASSITPRLTNDMTWEPKRRARD 359
QY 361 KOGVNAVMTSGHNDOSOMLOYDLVPTKYGITOGAKRFGYOVGSYKLAYSNDGEHM 420
DB 360 KOGVNAVMTSGHNDOSOMLOYDLVPTKYGITOGAKRFGYOVGSYKLAYSNDGEHM 419
QY 421 TVYDEKQKDKVFOGNFNDNTJHRKNYIDPPYARIRILRLMSYGRITLASLGCTEE 480
DB 420 TVYDEKQKDKVFOGNFNDNTJHRKNYIDPPYARIRILRLMSYGRITLASLGCTEE 479
QY 481 E 481
DB 1

```

```

DB 480 E 480

RESULT 4
US-09-237-981-10
Sequence 10, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quartermoun, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-10

Query Match          94.2%; Score 2519.5; DB 16; Length 480;
Best Local Similarity 94.2%; Pred. No. 3.5e-208;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCENGGICLPLGLAVSFCSECPDGFTPNCS 60
DB 1 MKHLVAMLLVGLSLGVPOFGKGDICPNPCENGGLSLGLADDSFSCDEPGEFAPNCS 60
QY 61 SVEVVASDEEPTSGAGPCTPMPCHNGSTCEISEAYRGDTFGLGYCKCPRGFGNCHQHN 120
DB 61 SVEVVASDEEPTSGAGPCTPMPCHNGSTCEISEAYRGDTFGLGYCKCPRGFGNCHQHN 120
QY 121 NECEVEPCKNGGICIDLVANYSCECPGEFGRNCOYKCSGGLIGGIIISNOGITASSSTH 180
DB 121 NECEVEPCKNGGICIDLVANYSCECPGEFGRNCOYKCSGGLIGGIIISNOGITASSSTH 180
QY 181 RALFGLQKWPYARLANKKGLINAMTAENDRKRWIOINLQRMKRYGVITOGAKRIGS 240
DB 181 RALFGLQKWPYARLANKKGLINAMTAENDRWP-WIOINLQRMKRYGVITOGAKRIGS 239
QY 241 PEYIKFKTIAYSNDGKTWAMTKVAGTNEDEVFRGNIDNNTPYANSFPPPIKAQYVALYQ 300
DB 240 PEYIKFKTIAYSNDGKTWAMTKVAGTNEDEVFRGNIDNNTPYANSFPPPIKAQYVALYQ 299

```

QY 301 VCRHCTLMELGCELSGCESEPLGMSGHIDYQITASSIFRILNMDMTWEPKARLD 360
:|||||
Db 300 ICRHCTLMELGCELSGCESEPLGMSGHIDYQITASSIFRILNMDMTWEPKARLD 359
QY 361 KQGVNANMTSGHNDOSQWLVQYDLVPTKYGTITGADPFHGVFGVSKYLA5NDGEHM 420
:|||||
Db 360 KQGVNANMTSGHNDOSQWLVQYDLVPTKYGTITGADPFHGVFGVSKYLA5NDGEHM 419
QY 421 TVYODEKORVDYFQGNFNDTHRKNNIDPPIYARHILRPMXYGRITLASELGCPE 480
:|||||
Db 420 MVHODEKORVDYFQGNFNDTHRKNNIDPPIYARHILRPMXYGRITLASELGCPE 479
QY 481 E 481
Db 480 E 480

RESULT 5
US-08-480-229B-21
: Sequence 21, Application US/08480229B
: GENERAL INFORMATION:
: APPLICANT: Quartermours, Thomas
: APPLICANT: Hogan, Bridgid
: APPLICANT: Snodgrass, H. Ralph
: APPLICANT: Zupancic, Thomas J.
: TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
: TITLE OF INVENTION: CELL LOCUS-1
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/480,229B
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Poissant, Brian M.
: REGISTRATION NUMBER: 28,462
: REFERENCE/DOCKET NUMBER: 8907-026
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: TELEX: 66141 Pennie
: INFORMATION FOR SEQ ID NO: 21:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 321 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
US-08-480-229B-21

Query Match 64.28; Score 1717; DB 8; Length 321;
Best Local Similarity 98.44; Pred. No. 2,3e-139;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 158 CCGPLGIEGGIISNOQITASTSTRALFGLQKWPYYARLNKGLINAMTAENDRWKRWI 217
:|||||
Db 1 CCGPLGIEGGIISNOQITASTSTRALFGLQKWPYYARLNKGLINAMTAENDRWKRWI 60
QY 218 QINLQKRMVTVITGAKRIGSEPEYIKFYKLA5NDGKTWAMYKVGITNEDWVFRGNID 277
:|||||

Db 61 QINLQKRMVTVITGAKRIGSEPEYIKFYKLA5NDGKTWAMYKVGITNEDWVFRGNID 120
QY 278 NNTPVANSFTPIKAOVRYLPVQVCRHCTLMELGCELSGCESEPLGMSGHIDYQIT 337
:|||||
Db 121 NNTPVANSFTPIKAOVRYLPVQVCRHCTLMELGCELSGCESEPLGMSGHIDYQIT 180
QY 338 ASSIFRITLMDMTWEPKARLDKQGVNANMTSGHNDOSQWLVQYDLVPTKYGTITGCA 397
:|||||
Db 181 ASSIFRITLMDMTWEPKARLDKQGVNANMTSGHNDOSQWLVQYDLVPTKYGTITGCA 240
QY 398 KDEGVQFVGSYKLA5NDGEHMTVYODEKORVDYFQGNFNDTHRKNNIDPPIYARH 457
:|||||
Db 241 KDXGHVQFVGSYKLA5NDGEHMTVYODEKORVDYFQGNFNDTHRKNNIDPPIYARH 300
QY 458 RILPMSWYGRITLASELGCCT 478
:|||||
Db 301 RILPMSWYGRITLASELGCCT 321

RESULT 6
US-08-659-235B-21
: Sequence 21, Application US/08659235B
: GENERAL INFORMATION:
: APPLICANT: Quartermours, Thomas
: APPLICANT: Hogan, Bridgid
: APPLICANT: Snodgrass, H. Ralph
: APPLICANT: Zupancic, Thomas J.
: TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
: TITLE OF INVENTION: CELL LOCUS-1
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/659,235B
: FILING DATE: 05-JUN-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Poissant, Brian M.
: REGISTRATION NUMBER: 28,462
: REFERENCE/DOCKET NUMBER: 8907-034
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: TELEX: 66141 Pennie
: INFORMATION FOR SEQ ID NO: 21:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 321 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
US-08-659-235B-21

Query Match 64.28; Score 1717; DB 10; Length 321;
Best Local Similarity 98.44; Pred. No. 2,3e-139;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 158 CCGPLGIEGGIISNOQITASTSTRALFGLQKWPYYARLNKGLINAMTAENDRWKRWI 217
:|||||
Db 1 CCGPLGIEGGIISNOQITASTSTRALFGLQKWPYYARLNKGLINAMTAENDRWKRWI 60
QY 218 QINLQKRMVTVITGAKRIGSEPEYIKFYKLA5NDGKTWAMYKVGITNEDWVFRGNID 277
:|||||

|||||
Db 61 QINLQRRKRVGVITOGAKRISPEYIKFYKIAVSNDGKTWAMYKVGNDMVRGNID 120
QY 278 NNTPYANSFPPPIKAQVRLYPQVCRHCTLRMELLCGCESEPLGKMSGHIDYOIT 337
Db 121 NNTPYANSFPPPIKAQVRLYPQVCRHCTLRMELLCGCESEPLGKMSGHIDYOIT 180
QY 338 ASSIFRRLNMDMFTWEPKARKLDKOGKVAMTSGHNDOSQMLQVLLVPTKVTGIITOGA 397
Db 181 ASSIFRRLNMDMFTWEPKARKLDKOGKVAMTSGHNDOSQMLQVLLVPTKVTGIITOGA 240
QY 398 KFGHVOFVGSYKLAYSNDGEHMTYVODEKORKDKVFGNFDNDRKKNVIDPPIYARHI 457
Db 241 KFGHVOFVGSYKLAYSNDGEHMTYVODEKORKDKVFGNFDNDRKKNVIDPPIYARHI 300
QY 458 RILPMSWYGRITLASSELLGCT 478
Db 301 RILPMSWYGRITLASSELLGCT 321
RESULT 7
US-09-237-981-21
: Sequence 21, Application US/09237981
: GENERAL INFORMATION:
: APPLICANT: Quettermous, Thomas
: APPLICANT: Hogan, Bridg
: APPLICANT: Snodgrass, H. Ralph
: APPLICANT: Zupancic, Thomas J.
: TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
: TITLE OF INVENTION: CELL LOCUS-1
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Penne & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/237,981
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/659,235
: FILING DATE: 05-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Poissant, Brian M.
: REGISTRATION NUMBER: 28,462
: REFERENCE/DOCKET NUMBER: 8907-0034-999
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 780-9090
: TELEFAX: (212) 869-8664/9741
: TELEX: 66141 Remnie
: INFORMATION FOR SEQ ID NO: 21:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 321 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-09-237-981-21
Query Match 64.2%; Score 1717; DB 16; Length 321;
Best Local Similarity 98.4%; Pred. No. 2,3e-139;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

|||||
Db 1 CSQPIGIGGIISNOQITASTSHRALFGLOKWPYARLNKKGILINAMTAENDRKNMI 60
QY 218 QINLQRRKRVGVITOGAKRISPEYIKFYKIAVSNDGKTWAMYKVGNDMVRGNID 277
Db 61 QINLQRRKRVGVITOGAKRISPEYIKFYKIAVSNDGKTWAMYKVGNDMVRGNID 120
QY 278 NNTPYANSFPPPIKAQVRLYPQVCRHCTLRMELLCGCESEPLGKMSGHIDYOIT 337
Db 121 NNTPYANSFPPPIKAQVRLYPQVCRHCTLRMELLCGCESEPLGKMSGHIDYOIT 180
QY 338 ASSIFRRLNMDMFTWEPKARKLDKOGKVAMTSGHNDOSQMLQVLLVPTKVTGIITOGA 397
Db 181 ASSIFRRLNMDMFTWEPKARKLDKOGKVAMTSGHNDOSQMLQVLLVPTKVTGIITOGA 240
QY 398 KFGHVOFVGSYKLAYSNDGEHMTYVODEKORKDKVFGNFDNDRKKNVIDPPIYARHI 457
Db 241 KFGHVOFVGSYKLAYSNDGEHMTYVODEKORKDKVFGNFDNDRKKNVIDPPIYARHI 300
QY 458 RILPMSWYGRITLASSELLGCT 478
Db 301 RILPMSWYGRITLASSELLGCT 321
RESULT 8
US-60-207-315-404
: Sequence 404, Application US/60207315
: GENERAL INFORMATION:
: APPLICANT: Beasley, Ellen
: TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
: FILE REFERENCE: C1000601
: CURRENT APPLICATION NUMBER: US/60/207,315
: CURRENT FILING DATE: 2000-05-30
: NUMBER OF SEQ ID NOS: 528
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 404
: LENGTH: 449
: TYPE: PRT
: ORGANISM: HUMAN
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(449)
: OTHER INFORMATION: Xaa = Any Amino Acid
: US-60-207-315-404
Query Match 60.4%; Score 1615; DB 23; Length 449;
Best Local Similarity 70.6%; Pred. No. 2.1e-130;
Matches 305; Conservative 2; Mismatches 7; Indels 118; Gaps 3;
QY 77 PCPBNCHNGGTCETSEAYRGDTFGYVCKCPGFNGJHCOH----- 118
Db 2 PCPBNCHNGGTCETSEAYRGDTFGYVCKCPGFNGJHCOH----- 61
QY 119 ----- 118
Db 62 KPLPLSLIDNAPRNSKKAARISTSSGRVPLSLTREPDMHEVGRKAMNTGLAMPYL 121
QY 119 -----NINECEYERCKNGGICDTLVANYSCECPGEPGRKCOYKCSGP 161
Db 122 RGGIICQLLRPECGAPNINECEYERCKNGGICDTLVANYSCECPGEPGRKCOYKCSGP 181
QY 162 LGIEGGIISNOQITASTSHRALFGLOKWPYARLNKKGILINAMTAENDRKN----- 214
Db 182 LGIEGGIISNOQITASTSHRALFGLOKWPYARLNKKGILINAMTAENDRKN----- 241
QY 215 ----RWIQLNLRKRVGVITOGAKRISPEYIKFYKIAVSNDGKTWAMYKVGNDM 270
Db 242 CHLGPMQDQINLQRRKRVGVITOGAKRISPEYIKFYKIAVSNDGKTWAMYKVGNDM 301
QY 271 VFGNIDNNTPYANSFPPPIKAQVRLYPQVCRHCTLRMELLCGCESEPLGKMSGH 330

```

Db 302 VERGIDNNTPYANSFPPKAYVRLPYOVCRRHCTLRMELLCGELSCSEPLGKMSGH 361
QY 331 IDVOYITASSFFRLNMDMTEPRKARLDKQGVNMTSGHNDOSQMLQ----- 380
Db 362 IDVOYITASSFFRLNMDMTEPRKARLDKQGVNMTSGHNDOSQMLQPTLISVEVSG 421
QY 381 --VDLLVPTKYT 390
Db 422 TALALMVPYAT 433

```

```

RESULT 9
US-60-207-315-523
; Sequence 323, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000601
; CURRENT APPLICATION NUMBER: US/60/207, 315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523
; LENGTH: 362
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-523

```

```

Query Match 52.6%; Score 1406.5; DB 23; Length 362;
Best Local Similarity 88.8%; Pred. No. 1.4e-112;
Matches 262; Conservative 3; Mismatches 7; Indels 23; Gaps 2;

QY 119 NINECEVEPCCKNGICTDLVANSCECPGEFMGRNOCYKCSPLGIGGIISNOQTASS 178
Db 52 DINECEVEPCCKNGICTDLVANSCECPGEFMGRNOCYKCSPLGIGGIISNOQTASS 111
QY 179 THRALFGLQKWPYYARLNKGLINAMTAENDRWK-----RWIQINLQRRMRY 227
Db 112 THRALFGLQKWPYYARLNKGLINAMTAENDRWPIQELRGCHLGPMDQINLQRRMRY 171
QY 228 TGVITOGAKRIGSPYIKFYKIAYSNDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSFT 287
Db 172 TGVITOGAKRIGSPYIKFYKIAYSNDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSFT 231
QY 288 PIKAOYVRLPYOVCRRHCTLRMELLCGELSCSEPLGKMSGHIDVOYITASSIFRLNM 347
Db 232 PIKAOYVRLPYOVCRRHCTLRMELLCGELSCSEPLGKMSGHIDVOYITASSIFRLNM 291
QY 348 DMFTWEPKARLDKQGVNMTSGHNDOSQMLQ-----VDLLVPTKYT 390
Db 292 DMFTWEPKARLDKQGVNMTSGHNDOSQMLQPTLISVEVSGTALALMVPYAT 346

```

```

RESULT 10
US-60-230-435-1519
; Sequence 1519, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230, 435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1519
; LENGTH: 311
; TYPE: PRT

```

```

; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(311)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-230-435-1519

```

```

Query Match 52.5%; Score 1405.5; DB 23; Length 311;
Best Local Similarity 89.1%; Pred. No. 1.4e-112;
Matches 262; Conservative 2; Mismatches 7; Indels 23; Gaps 2;

QY 120 INCEVEPCCKNGICTDLVANSCECPGEFMGRNOCYKCSPLGIGGIISNOQTASS 179
Db 2 INCEVEPCCKNGICTDLVANSCECPGEFMGRNOCYKCSPLGIGGIISNOQTASS 61
QY 180 HRALFGLQKWPYYARLNKGLINAMTAENDRWK-----RWIQINLQRRMRY 228
Db 62 HRALFGLQKWPYYARLNKGLINAMTAENDRWPIQELRGCHLGPMDQINLQRRMRY 121
QY 229 GVTIOGAKRIGSPYIKFYKIAYSNDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSFT 288
Db 122 GVTIOGAKRIGSPYIKFYKIAYSNDGKTWAMYKVGTDNEDVFRGNIDNNTPYANSFT 181
QY 289 PIKAOYVRLPYOVCRRHCTLRMELLCGELSCSEPLGKMSGHIDVOYITASSIFRLNM 348
Db 182 PIKAOYVRLPYOVCRRHCTLRMELLCGELSCSEPLGKMSGHIDVOYITASSIFRLNM 241
QY 349 MFTWEPKARLDKQGVNMTSGHNDOSQMLQ-----VDLLVPTKYT 390
Db 242 MFTWEPKARLDKQGVNMTSGHNDOSQMLQPTLISVEVSGTALALMVPYAT 295

```

```

RESULT 11
US-09-582-340-4
; Sequence 4, Application US/09582340
; GENERAL INFORMATION:
; APPLICANT: INSERM
; APPLICANT: INSTITUT CURIE
; APPLICANT: CNRS
; TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants
; FILE REFERENCE: Lactadherin
; CURRENT APPLICATION NUMBER: US/09/582, 340
; CURRENT FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 463
; TYPE: PRT
; ORGANISM: mammalian
US-09-582-340-4

```

```

Query Match 43.2%; Score 1156; DB 19; Length 463;
Best Local Similarity 48.5%; Pred. No. 6.7e-91;
Matches 223; Conservative 72; Mismatches 142; Indels 22; Gaps 8;

QY 23 GDICDPNCPENGICLPGLAVSGSCPCPDGFTDPCSSVVEVASDEEPTAGCTPNP 82
Db 25 GDICDPNCPENGICLPGLAVSGSCPCPDGFTDPCSSVVEVASDEEPTAGCTPNP 72
QY 83 CHNGGTCEIS-EAYRGDTFTGYVCKPRGFNGIHQHNINECEVEPCCKNGICTDLVANY 141
Db 73 CYNDKCKLVTLDTORGDIETFEYICQPVYSGIHCEETVNYVLD--GEVMTVAVPNT 129
QY 142 SCECGEE--GRRNCQYKCSPLGIGGIISNOQTASSSTRALFGLQKWPYYARLNK 199
Db 130 AVPTAPRPDLSSNNLASRSTQLGMGGAIDDSQASVYWMGFMGLQWGPRLRYRT 189
QY 200 GLINAMTAENDRWKRWIQINLQRRMRYGVTIOGAKRIGSPYIKFYKIAYSNDGKTWA 259
Db 190 GLINAMTAENDRWKRWIQINLQRRMRYGVTIOGAKRIGSPYIKFYKIAYSNDGKTWA 248

```



```

FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(185)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-230-435-1806

```

```

Query Match
Best Local Similarity 93.3%; Score 1036.5; DB 23; Length 185;
Matches 182; Conservative 0; Mismatches 2; Indels 11; Gaps 2;

```

```

QY 24 DICPNCENGICLPGLANGSFCPCPDGTFDPCSSVVEADEEPTSAEPCIPNC 83
DB 2 DICDPNCEMNGICLPGLANGSFCPCPDGTFDPCSSVVEADEEPTSAEPCIPNC 51
QY 84 HNGGCEISEAYRGDTFPIGVCKPRGFNGICQHINECEVEPCCKNGICTDLVANYSC 143
DB 52 HNGGCEISEAYRGDTFPIGVCKPRGFNGICQHINECEVEPCCKNGICTDLVANYSC 111
QY 144 ECPGPFMGRNCOYKCSGPIEGGITSNOQITASSSTHRALFGLQKWPYARLKKGLIN 203
DB 112 ECPGPFMGRNCOYKCSGPIEGGITSNOQITASSSTHRALFGLQKWPYARLKKGLIN 171
QY 204 AWTAAENDRMRKWTQ 218
DB 172 AWTAAENDRMRKWTQ 185

```

```

RESULT 15
US-09-582-340-2

```

```

; Sequence 2, Application US/09582340
; GENERAL INFORMATION:

```

```

; APPLICANT: INSERM
; APPLICANT: INSTITUT CURIE
; APPLICANT: CNRS
; TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants
; FILE REFERENCE: Lactadherin
; CURRENT APPLICATION NUMBER: US/09/582,340
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-582-340-2

```

```

Query Match
Best Local Similarity 48.8%; Score 1020.5; DB 19; Length 387;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;

```

```

QY 78 CTPNCHNGGTC-EISFAYRGDTFPIGVCKPRGFNGICQHINECEVEPCCKNGICTD 136
DB 27 CSKNPCNHGGICETISQEVGRGVDPSTCTCLKGAGNH----- 65
QY 137 LVANYSCBCEPGEFMGRNCOYKCSGPIEGGITSNOQITASSSTHRALFGLQKWPYARL 196
DB 66 -----CEFKCEVEPLGMENGINANSQIASSSVRTFGLQHWPELRL 108
QY 197 NKKGLINAWTAENDRMRKWTQINLQKRWGVITQCAKRGSPSEYIKFYKIAYSNDGK 256
DB 109 NRAGVNVNMTPTSSND-NPFIQVNLRRMWTQVYTGASRLASHHEYLAKFVAYSLNGH 167
QY 257 TW-AMTKYKGTNEDMVRGINDNNTPYANSETPPIKAQYVRLYPQVCRHCTLMELLGC 315
DB 168 EEDFLHDVNRKHKKEFV--GWMNRNAVHVLFEFPEAQYVRLYPTSCHTACTLRFELLGC 225
QY 316 ELSGCSPEPLGKSGHIDYQITASSIFRTLINMDMFWEPKRLDKQGVNANMTSGHNDQ 375
DB 226 ELNGCANPLGLKNSIDPKQITASSSVKWTGHLFPMNPSYARLDKQGNFANWVAGSYGN 285

```

```

QY 376 SQWLOYDLIVPTKNGITTOGAKDEGHVQFVGSYKLYASNDGHEHMTYODEKQKDKYFQ 435
DB 286 DQWLOYDLSSKEVETGITTQAKNFGSVQFYASYKAYSNDSANMTETQDPTGSSKIFP 345
QY 436 GNFDNDTHRKNVIDPPIYARHRIPLPWSWGRITLASELLGC 477
DB 346 GNMNDNHRKKNLFEETPLARYRILPVAMHNRIALRLLELLGC 387

```

```

Search completed: May 23, 2001, 08:39:40
Job time: 131 sec

```


Wed May 23 08:50:27 2001

us-09-237-981-14_copy_33_513.clo.rapn

Page 4

QY 167 -----GIISNOOITASSTH 180
DB 329 QEGWGHGCKNKRYEASLIH 347

RESULT 9

US-09-423-844-4
; Sequence 4, Application US/09423844
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Chen, Jian
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth
; APPLICANT: Penicka, Diane
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P1618P1E
; CURRENT APPLICATION NUMBER: US/09/423,844
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-844-4

Query Match

Best Local Similarity 25.6%, Score 182; DB 5; Length 379;
Matches 51; Conservative 16; Mismatches 52; Indels 80; Gaps 10;

QY 28 PNCENGICLPGLAVSFSCEPDGFTDPNCSSVVEVASEDEEPTISAGCTPNPCHNG 87
DB 183 PGCRNGGFCNERRI-----CECPDGFHGHCKAL-----CTPR-CMNG 222
QY 88 TCISEAYRGDTFIVGCKPRGNGIHC-----Q 117
DB 223 LC-VTPGF-----CICPFGFYGVCDKANCSTTCFNGGTCFYPKCICPGLSEQ 272
QY 118 HINICEVEPCKNGICTDLVANYSCGCEPFGMGRNCOYK-CGSPGLIGS----- 166
DB 273 CEISKCP-QPCRNGKC---IGSKCKCKSGYGDLSKPYCEPGCGAHGTCHHPNKCQC 328
QY 167 -----GIISNOOITASSTH 180
DB 329 QEGWGHGCKNKRYEASLIH 347

RESULT 10

US-09-403-296A-4
; Sequence 4, Application US/09403296A
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
; FILE REFERENCE: P1624R2E
; CURRENT APPLICATION NUMBER: US/09/403,296A
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/100,858
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/19437
; PRIOR FILING DATE: 1998-09-17

; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-403-296A-4

Query Match

Best Local Similarity 25.6%, Score 182; DB 5; Length 379;
Matches 51; Conservative 16; Mismatches 52; Indels 80; Gaps 10;

QY 28 PNCENGICLPGLAVSFSCEPDGFTDPNCSSVVEVASEDEEPTISAGCTPNPCHNG 87
DB 183 PGCRNGGFCNERRI-----CECPDGFHGHCKAL-----CTPR-CMNG 222
QY 88 TCISEAYRGDTFIVGCKPRGNGIHC-----Q 117
DB 223 LC-VTPGF-----CICPFGFYGVCDKANCSTTCFNGGTCFYPKCICPGLSEQ 272
QY 118 HINICEVEPCKNGICTDLVANYSCGCEPFGMGRNCOYK-CGSPGLIGS----- 166
DB 273 CEISKCP-QPCRNGKC---IGSKCKCKSGYGDLSKPYCEPGCGAHGTCHHPNKCQC 328
QY 167 -----GIISNOOITASSTH 180
DB 329 QEGWGHGCKNKRYEASLIH 347

RESULT 11

US-09-380-139A-4
; Sequence 4, Application US/09380139A
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC. et al.
; APPLICANT: Chen, Jian
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Penicka, Diane
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: 10466-04
; CURRENT APPLICATION NUMBER: US/09/380,139A
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: PCT/US 98/19330
; PRIOR FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-380-139A-4

Query Match

Best Local Similarity 25.6%, Score 182; DB 5; Length 379;
Matches 51; Conservative 16; Mismatches 52; Indels 80; Gaps 10;

QY 28 PNCENGICLPGLAVSFSCEPDGFTDPNCSSVVEVASEDEEPTISAGCTPNPCHNG 87
DB 183 PGCRNGGFCNERRI-----CECPDGFHGHCKAL-----CTPR-CMNG 222
QY 88 TCISEAYRGDTFIVGCKPRGNGIHC-----Q 117
DB 223 LC-VTPGF-----CICPFGFYGVCDKANCSTTCFNGGTCFYPKCICPGLSEQ 272
QY 118 HINICEVEPCKNGICTDLVANYSCGCEPFGMGRNCOYK-CGSPGLIGS----- 166
DB 273 CEISKCP-QPCRNGKC---IGSKCKCKSGYGDLSKPYCEPGCGAHGTCHHPNKCQC 328
QY 167 -----GIISNOOITASSTH 180

Wed May 23 08:50:25 2001

US-09-237-981-14_COPY_33_513.c10.rat

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 08:36:49 ; Search time 12.57 Seconds
(without alignments)
735.117 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513

Perfect score: 2675
Sequence: 1 MKRSVAVMLVGLSGVPOF.....MSWYKHTLASFLGCTEE 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 175427

Minimum DB seq length: 0

Maximum DB seq length: 481

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2519.5	94.2	480	2	US-08-480-229C-10
2	2519.5	94.2	480	2	US-08-659-235C-10
3	1717	64.2	321	2	US-08-480-229C-21
4	1717	64.2	321	2	US-08-659-235C-21
5	1144	42.8	463	2	US-08-480-229C-29
6	1133.5	42.4	221	2	US-08-480-229C-29
7	1133.5	42.4	221	2	US-08-659-235C-29
8	1120.5	41.9	462	2	US-08-480-229C-29
9	1020.5	35.1	387	2	US-08-480-229C-20
10	948	35.4	320	2	US-08-659-235C-20
11	643	24.0	217	1	US-07-607-538C-3
12	643	24.0	217	1	US-08-162-402B-3
13	643	24.0	218	1	US-07-607-538C-2
14	643	24.0	218	1	US-08-162-402B-2
15	501	18.9	157	2	US-08-162-402B-13
16	498	18.6	159	2	US-08-162-402B-12
17	498	18.6	159	2	US-08-162-402B-12
18	498	18.6	159	2	US-08-162-402B-12
19	498	18.6	159	2	US-08-162-402B-12
20	494	18.5	218	1	US-07-607-538C-4
21	472.5	17.8	218	1	US-08-162-402B-5
22	472.5	17.8	218	1	US-08-162-402B-5
23	414	15.2	160	2	US-08-162-402B-11
24	387	14.5	85	2	US-08-457-135C-1
25	378	14.5	85	2	US-08-457-135C-1
26	378	14.5	160	2	US-08-162-402B-14
27	368	13.8	109	1	US-08-111-939-25

28	344.5	12.9	160	2	US-08-162-402B-15
29	305.5	11.4	154	2	US-08-162-402B-17
30	305	11.4	150	1	US-08-111-939-24
31	293.5	10.6	159	2	US-08-162-402B-16
32	284.5	10.6	57	2	US-08-480-229C-23
33	284.5	10.6	57	2	US-08-659-235C-23
34	264	9.9	43	2	US-08-480-229C-24
35	264	9.9	43	2	US-08-659-235C-24
36	257	9.6	111	4	US-08-111-939-23
37	256	9.6	157	4	US-08-872-855-6
38	252	9.4	85	2	US-08-480-229C-2
39	245	9.4	85	2	US-08-659-235C-2
40	245	9.4	85	2	US-08-457-135C-2
41	245	9.2	383	1	US-08-457-135C-2
42	244.5	9.1	385	1	US-08-457-135C-1
43	244.5	9.1	385	1	US-08-457-135C-1
44	232.5	8.7	109	1	US-08-111-939-23
45	227	8.5	161	2	US-08-162-402B-19

ALIGNMENTS

Result 1
US-08-480-229C-10
Sequence 10 Application US/08480229C
Patent No 5874562
GENERAL INFORMATION:
APPLICANT: Quaternous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penne & Edmunds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentio Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELEPHONE: (212) 780-9090
TELEFAX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229C-10

Query Match 94.2% Score 2519.5 DB 2: Length 480:
Best Local Similarity 94.2% Pred. No. 1.2e+197:
Matches 453; Conservative 9; Mismatches 18; Indels 1:
Gaps 1:
1 MKRSVAVMLVGLSGVPOFGKGDICDPNFCENGICLRLAVGSFSCDCHDTPNCS 60

```

Db 1 MRLVLAAMLVLGSLGVPOFGKGDICNPENCGICLSGLADDSFSCDEPESGAPNCS 60
QY 1 SVVEVADSEDEEPTPSAGPCIPNCPHNGGTCEISEAYRGDTFYGVCCKPRGPNHICQNI 120
Db 61 SVVEVADSEDEEPTPSAGPCIPNCPHNGGTCEISEAYRGDTFYGVCCKPRGPNHICQNI 120
QY 121 NCEVEPCKNGIGCTDLVANSCECPGEMGRNCOYKCSGPLGIEGIIISNOQITASSTH 180
Db 121 NCEVEPCKNGIGCTDLVANSCECPGEMGRNCOYKCSGPLGIEGIIISNOQITASSTH 180
QY 181 RALFELQKWPYYTARLNKGLINAMTAENDRMKRMIOINLQRMKRYGVITOGARRIGS 240
Db 181 RALFELQKWPYYTARLNKGLINAMTAENDRMKRMIOINLQRMKRYGVITOGARRIGS 240
QY 241 PEYIKFKIAYSDGKTAMKVKGTNEDEVFNGNIDNNTPIANSFTPIKAQYVRLYPO 300
Db 241 PEYIKFKIAYSDGKTAMKVKGTNEDEVFNGNIDNNTPIANSFTPIKAQYVRLYPO 300
QY 301 VCRHCTLRMELLGCELSGSEPLGKSGHIOYOTASSITRILNMDTWEPRARLD 360
Db 301 VCRHCTLRMELLGCELSGSEPLGKSGHIOYOTASSITRILNMDTWEPRARLD 360
QY 361 KQGVNMTSGHNDOSQMLVDLVPKVGITTOGAKPFGHVPVGSYKLAYSNGEHW 420
Db 361 KQGVNMTSGHNDOSQMLVDLVPKVGITTOGAKPFGHVPVGSYKLAYSNGEHW 420
QY 421 PYVDEORQKDKYFQGNFNDHNRNYIDPPIYARIRILPMSYGRITLASLLOCTEE 480
Db 421 PYVDEORQKDKYFQGNFNDHNRNYIDPPIYARIRILPMSYGRITLASLLOCTEE 480
QY 481 E 481
Db 480 E 480

```

```

RESULT 2
US-08-659-235C-10
; Sequence 10, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Queternous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Polsant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8866/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235C-10

```

```

Query Match 94.2%; Score 2519.5; DB 2; Length 480;
Best Local Similarity 94.2%; Pred. No. 1,2e-197;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

```

```

QY 1 MRSVAVMLVLGSLGVPOFGKGDICDPENCGICLPLGLAVSGSFSCDEPESGAPNCS 60
Db 1 MRLVLAAMLVLGSLGVPOFGKGDICNPENCGICLSGLADDSFSCDEPESGAPNCS 60
QY 61 SVVEVADSEDEEPTPSAGPCIPNCPHNGGTCEISEAYRGDTFYGVCCKPRGPNHICQNI 120
Db 61 SVVEVADSEDEEPTPSAGPCIPNCPHNGGTCEISEAYRGDTFYGVCCKPRGPNHICQNI 120
QY 121 NCEVEPCKNGIGCTDLVANSCECPGEMGRNCOYKCSGPLGIEGIIISNOQITASSTH 180
Db 121 NCEVEPCKNGIGCTDLVANSCECPGEMGRNCOYKCSGPLGIEGIIISNOQITASSTH 180
QY 181 RALFELQKWPYYTARLNKGLINAMTAENDRMKRMIOINLQRMKRYGVITOGARRIGS 240
Db 181 RALFELQKWPYYTARLNKGLINAMTAENDRMKRMIOINLQRMKRYGVITOGARRIGS 240
QY 241 PEYIKFKIAYSDGKTAMKVKGTNEDEVFNGNIDNNTPIANSFTPIKAQYVRLYPO 300
Db 241 PEYIKFKIAYSDGKTAMKVKGTNEDEVFNGNIDNNTPIANSFTPIKAQYVRLYPO 300
QY 301 VCRHCTLRMELLGCELSGSEPLGKSGHIOYOTASSITRILNMDTWEPRARLD 360
Db 301 VCRHCTLRMELLGCELSGSEPLGKSGHIOYOTASSITRILNMDTWEPRARLD 360
QY 361 KQGVNMTSGHNDOSQMLVDLVPKVGITTOGAKPFGHVPVGSYKLAYSNGEHW 420
Db 361 KQGVNMTSGHNDOSQMLVDLVPKVGITTOGAKPFGHVPVGSYKLAYSNGEHW 420
QY 421 PYVDEORQKDKYFQGNFNDHNRNYIDPPIYARIRILPMSYGRITLASLLOCTEE 480
Db 421 PYVDEORQKDKYFQGNFNDHNRNYIDPPIYARIRILPMSYGRITLASLLOCTEE 480
QY 481 E 481
Db 480 E 480

```

```

RESULT 3
US-08-480-229C-21
; Sequence 21, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Queternous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

Wed May 23 08:50:25 2001

us-09-237-981-14_copy_33_513.c10.ra1

Page 3

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-229C-21

```

```

Query Match      64.2%  Score 1717; DB 2; Length 321;
Best Local Similarity 98.4%  Pred. No. 1.8e-132;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 158 CSGPLGEGGIISNOQITVASTTHRALFGLOKWPYARLNKKGILINMTAENDRWKRWI 217
    1 CSGPLGEGGIISNOQITVASTTHRALFGLOKWPYARLNKKGILINMTAENDRWKRWI 60
DB 1 QINQRMKRVTVITQOGAKRIGSPYEIKFYKLIAYSDGKTAMKVKVGTNEDVFRGNID 277
QY 218 QINQRMKRVTVITQOGAKRIGSPYEIKFYKLIAYSDGKTAMKVKVGTNEDVFRGNID 277
    61 QINQRMKRVTVITQOGAKRIGSPYEIKFYKLIAYSDGKTAMKVKVGTNEDVFRGNID 120
DB 278 NNPVANSFTPIPKAOYRLPYOCRRHCTLRMELIGCELSGCESEPLGMSGHIIDYQIT 337
QY 121 NNPVANSFTPIPKAOYRLPYOCRRHCTLRMELIGCELSGCESEPLGMSGHIIDYQIT 180
DB 338 ASSIFRLNMDMFTPEPKARLDKOGVNMVSGHNDOSOMLYQDLVLPKRYGIIITGGA 397
QY 181 ASSIFRLNMDMFTPEPKARLDKOGVNMVSGHNDOSOMLYQDLVLPKRYGIIITGGA 240
DB 398 KRGHVOFVGSYKLAISNDGEHMTVYODEKOROKVFGCNFDNTHRKNVTDPPITARIH 457
QY 241 KRGHVOFVGSYKLAISNDGEHMTVYODEKOROKVFGCNFDNTHRKNVTDPPITARIH 300
DB 458 RILPMSWYGRITLASELLGCT 478
QY 301 RILPMSWYGRITLASELLGCT 321
DB

```

```

RESULT 4
US-08-659-235C-21
; Sequence 21, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-659-235C-21

```

```

Query Match      64.2%  Score 1717; DB 2; Length 321;
Best Local Similarity 98.4%  Pred. No. 1.8e-132;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 158 CSGPLGEGGIISNOQITVASTTHRALFGLOKWPYARLNKKGILINMTAENDRWKRWI 217
    1 CSGPLGEGGIISNOQITVASTTHRALFGLOKWPYARLNKKGILINMTAENDRWKRWI 60
DB 1 QINQRMKRVTVITQOGAKRIGSPYEIKFYKLIAYSDGKTAMKVKVGTNEDVFRGNID 277
QY 218 QINQRMKRVTVITQOGAKRIGSPYEIKFYKLIAYSDGKTAMKVKVGTNEDVFRGNID 277
    61 QINQRMKRVTVITQOGAKRIGSPYEIKFYKLIAYSDGKTAMKVKVGTNEDVFRGNID 120
DB 278 NNPVANSFTPIPKAOYRLPYOCRRHCTLRMELIGCELSGCESEPLGMSGHIIDYQIT 337
QY 121 NNPVANSFTPIPKAOYRLPYOCRRHCTLRMELIGCELSGCESEPLGMSGHIIDYQIT 180
DB 338 ASSIFRLNMDMFTPEPKARLDKOGVNMVSGHNDOSOMLYQDLVLPKRYGIIITGGA 397
QY 181 ASSIFRLNMDMFTPEPKARLDKOGVNMVSGHNDOSOMLYQDLVLPKRYGIIITGGA 240
DB 398 KRGHVOFVGSYKLAISNDGEHMTVYODEKOROKVFGCNFDNTHRKNVTDPPITARIH 457
QY 241 KRGHVOFVGSYKLAISNDGEHMTVYODEKOROKVFGCNFDNTHRKNVTDPPITARIH 300
DB 458 RILPMSWYGRITLASELLGCT 478
QY 301 RILPMSWYGRITLASELLGCT 321
DB

```

```

RESULT 5
US-08-162-402B-9
; Sequence 9, Application US/08162402B
; Patent No. 5972337
; GENERAL INFORMATION:
; APPLICANT: CERIANI, ROBERTO L.
; APPLICANT: PETERSON, JERRY A.
; APPLICANT: LABOCCA, DAVID J.
; TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
; TITLE OF INVENTION: GLOBULE (HMFg) ANTIGEN, FRAGMENTS & FUSION PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder & Poplawski
; STREET: 444 South Flower St., 19th Floor
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

```

```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162.402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-9

```

```

Query Match          42.8% Score 1144; DB 2; Length 463;
Best Local Similarity 48.0%; Pred. No. 1,4e-85;
Matches 220; Conservative 72; Mismatches 144; Indels 22; Gaps 8;

```

```

QY 23 GDICDPNCEGICLPLAVGSCCECPDGFDPNCSVAVASDEDEPTSGPCPP 82
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 25 GDFDSSLCLNGTCLTG-QDNDIYCLCPSTFTGLVGN-TERGPCSPNP 72
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 83 CHNGTGEIS-EAYRGDTFGLVCKCPRGNGIHCOHNTNECEPRCKNGICDIDY 141
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 73 CYNKAKLVLPDORDIFETICCPVSGIHCEETNYNLD--GEMPTAVPMT 129
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 142 SCEPRGEF--MGRNCOYKCSGPIEGGITSNNOITASTHRALEGLQKWPYARLNK 199
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 130 AVPRATPPLSNLAKRSQTGLGSGALADSDISNIVYGFMDKMGPELARLYRT 189
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 200 GLINAMTAENDMKRWIOLINQKMKRTGYITQGAKRIGSPYIKFKIAYSNDQKTA 259
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 190 GIVNMAHSAKVDLP-WIQVNLKMKRYSGVMTQGSRAGRALYKTFKVAISLDGKFE 248
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 260 MYKAGTNEDMVTRGNIDNTPYANSFTPRKAQVYVLYPYOYCRHCTLRRELLGCELSG 319
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 249 FLODE-SGGRKFLGMLDNNLSLKVMMPTLEAYLRTPVSCHGCTLRRELLGCELSG 307
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 320 CSEPLGKSGHIQDYITASSIFTLNDMEFTWEPKARLDKQGVNANMTSGHNDOSQWL 379
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 308 CLEPLGKKNNTIPDSQMSASSSYKTWMLRAFGWYPHIGRLDNOGINAMTQSSAKRWL 367
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 380 QVDLLVPTKVTGLITQGAKEGHVQFVGSYKRLAYSNDGEMHTVYODEKORKDKVFOGND 439
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 368 QVDLGTORQVYGLITQGAKEGHVQFVGSYKRLAYSNDGEMHTVYODEKORKDKVFOGND 425
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 440 NDTIRKKNVIDPPIYARHRIILPWSWYGRITLASELLGC 477
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 426 NNSHKKNIFKPRMAKRVKVLVPSWHRITLRLLELLGC 463
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

RESULT 6
US-08-480-229C-29
Sequence 29, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph

```

```

APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-480-229C-29

```

```

Query Match          42.4% Score 1133.5; DB 2; Length 221;
Best Local Similarity 90.0%; Pred. No. 4e-85;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

```

```

QY 1 MKRSVAVMLVLSLGVQFGKGDICDPNCEGICLPLAVGSCCECPDGFDPNCS 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1 MKHLVAMLLVLSLGVQFGKGDICDPNCEGICLPLAVGSCCECPDGFDPNCS 60
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 SYVEVASDEEPTSGAGPCTPNPCHNGTGEISEAYRGDTFTGYVCKCPRGFNGIHCOHNT 120
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 61 SYVEVASDEEPTSGAGPCTPNPCHNGTGEISEAYRGDTFTGYVCKCPRGFNGIHCOHNT 120
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 NECEVEPRCKNGICDIDVANSYSCCEPRGEMGRNCOYKCSGPIEGGITSNNOITASSSTH 180
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 121 NECEVEPRCKNGICDIDVANSYSCCEPRGEMGRNCOYKCSGPIEGGITSNNOITASSSTH 180
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 RALFGLQKWPYARLNKKGILINAMTAENDMKRWIOLINL 221
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 181 RALFGLQKWPYARLNKKGILINAMTAENDMKRWIOLINL 220
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

RESULT 7
US-08-659-235C-29
Sequence 29, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEO ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-659-235C-29

Query Match 42.4%; Score 1133.5; DB 2; Length 221;
Best Local Similarity 90.0%; Pred No 4e-65; Indels 1; Gaps 1;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;
QY 1 MKRSVAVLLVGLSLGVRKRGICDPNCPNCGGICPLGLAVSFSCDECPDFTDPNKS 60
DB 1 MKHLYAVMLVGLSLGVRKRGICDPNCPNCGGICPLGLAVSFSCDECPDFTDPNKS 60
QY 61 SVYVAVADEEPTSAQCTPNPCNGGTCEISAVRGDPTFGYVCKCPGFGNGICQNIHCEVEP 120
DB 61 SVYVAVADEEPTSAQCTPNPCNGGTCEISAVRGDPTFGYVCKCPGFGNGICQNIHCEVEP 120
QY 121 NECEVECKRGICIDLVANYSCEGCEPEMRCQYKCSGPIGEGGISNOGITASSSTH 180
DB 121 NECAEPCRGICIDLVANYSCEGCEPEMRCQYKCSGPIGEGGISNOGITASSSTH 180
QY 181 RALGGLQKWPYVYARLNKGLINMTAAENDRKRPIQINL 221
DB 181 RALGGLQKWPYVYARLNKGLINMTAAENDRKRPIQINL 221

RESULT 9
US-08-162-402B-8
Sequence 8 Application US/08162402B
Patent No 5972337
GENERAL INFORMATION:
APPLICANT: CERIANI, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
APPLICANT: IAROSCA, DAVID J.
TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
TITLE OF INVENTION: GLOBULE (HMF) ANTIGEN FRAGMENTS & FUSION PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Petley, Schroeder & Poplawski
STREET: 444 South Flower St., 15th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEO ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-8

Query Match 41.9%; Score 1120.5; DB 2; Length 465;
Best Local Similarity 47.6%; Pred No 1.2e-83;
Matches 225; Conservative 63; Mismatches 154; Indels 31; Gaps 9;
QY 9 LTVGLSLGVRKRGICDPNCPNCGGICPLGLAVSFSCDECPDFTDPNKS SVYVAVASD 68
DB 20 LTVGLSLGVRKRGICDPNCPNCGGICPLGLAVSFSCDECPDFTDPNKS SVYVAVASD 68
QY 69 EEPFSAQCTPNPCNGGTCEISAVRGDPTFGYVCKCPGFGNGICQNIHCEVEP 127
DB 63 EEPFSAQCTPNPCNGGTCEISAVRGDPTFGYVCKCPGFGNGICQNIHCEVEP 127
QY 128 CNGSICIDLVANYSCEGCEPEMRCQYKCSGPIGEGGISNOGITASSSTH 185
DB 118 CNGSICIDLVANYSCEGCEPEMRCQYKCSGPIGEGGISNOGITASSSTH 185
QY 186 LKRWVYVYARLNKGLINMTAAENDRKRPIQINLORAKRVGTITGAKRISPEYIK 245
DB 176 LKRWVYVYARLNKGLINMTAAENDRKRPIQINLORAKRVGTITGAKRISPEYIK 245
QY 246 FKRIYVNDGKTW-AMTKVKGTEDEVRGNIQNPYANSFTPIKAQVRYLPQVCR 304
DB 235 FKRIYVNDGKTW-AMTKVKGTEDEVRGNIQNPYANSFTPIKAQVRYLPQVCR 304
QY 305 HCTLMELLGELSGCEPEMRCQYKCSGPIGEGGISNOGITASSSTH 364
DB 293 HCTLMELLGELSGCEPEMRCQYKCSGPIGEGGISNOGITASSSTH 364
QY 365 VNAVATSGHNDQGLVLDVPTKVTGTTGAKRFGHVOFGSKYKLVASNDGSHMTVQ 424
DB 353 VNAVATSGHNDQGLVLDVPTKVTGTTGAKRFGHVOFGSKYKLVASNDGSHMTVQ 424
QY 425 DKRQKDVQGNPDNTHKKNVLDPIYARHRIIDPMYSYGITLASELLIG 477
DB 413 DKRQKDVQGNPDNTHKKNVLDPIYARHRIIDPMYSYGITLASELLIG 477

RESULT 9
US-08-162-402B-6
Sequence 6 Application US/08162402B
Patent No 5972337
GENERAL INFORMATION:
APPLICANT: CERIANI, ROBERTO L.

Wed May 23 08:50:25 2001

us-09-237-981-14_copy_33_513.clo.ra1

Page 6

APPLICANT: PETERSON, JERRY A.
 APPLICANT: LAROCCA, DAVID J.
 TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
 TITLE OF INVENTION: GLOBULE (HMG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Pretty, Schroeder & Poplawski
 STREET: 444 South Flower St., 19th Floor
 CITY: Los Angeles
 STATE: CA
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/162,402B
 FILING DATE: 03-DEC-1993
 CLASSIFICATION: 435
 APPLICATION DATA:
 PRIOR APPLICATION NUMBER:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Amzel, Viviana
 REGISTRATION NUMBER: 30,930
 REFERENCE/DOCKET NUMBER: P66 38215
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 213-622-7700
 TELEFAX: 213-489-4210
 TELEFAX:
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 387 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 PS-08-162-402B-6

Query Match	Similarity	38 13	Score	1020.5	DB: 2	Length	387
Best Local	Similarity	48.8%	Pred. No.	1.3e-75			
Matches	196	Conservative	55	Mismatches	108	Indels	43
							Gaps
QY	78	CGPAPGNGGTG-EISEKVRDPFIQYCKCRBFGNHCQNIWCEVEFPNGSICTD	136				
DB	27	CGKNCKNGSLCEIISQYRBDYFPSTITCTCKIAGNH-----	65				
QY	137	LVANTSECPGEGEPMRCQYKSGPLGIEGGISNOQTASSTHRALPGIQKRYVABL	196				
DB	66	-----CEKCYEPLGEMNGNINSSAISSAVRYTEGLQIHWPELLRL	108				
QY	197	NKGLINAMTAENDRMKRWIOTINLORKRWYGVITOGAKRIGSPYIKFYKIAYSNDGK	256				
DB	109	NNAEGVAMTPRESSND-NPMLOYNLRRRWYTGVTGASRLASHYELKFKFAVSLNGH	167				
QY	257	TW-AMRYKVGKTEDWYFEGNIDNNTPYANSFPIKAQVRYRLPYQVCRHRCITLMELLGC	315				
DB	168	EEDFIEHOVNNKHKEEV--GNMKNNAVHNVLEFETVEEAQVRYRLPYSCITACTLREFLELLGC	225				
QY	316	ELSGCSEPELMSGSHIODYQITASSIFRTLNMDFTWEPERKARLDQOGKNAATSGSHDQ	375				
DB	226	ELNGCAPRLDGLNNNSIPDKQITASSSYKTKWGLHFSWNPYSARLDQGNINAAVASQGN	285				
QY	376	SGMLQVLDLVPKRVGIIITOGADPFHQVQFSGVSKLAYSNDENHTVYQDEKORQKQYFQ	435				
DB	286	DDMLQVLDLSSKKEVGIITOGARNPFSSQVQFVASYKAYVNSDSANNTETQDPPRTGSSKIFP	345				
QY	436	GNFNDNTHRKANYLDPPYARHRIKLWMSYGHILASSELLGC	477				
DB	346	GNMDSHSHKKNLFEPIRLARVYLLDVAANHNRLRLRELLGC	387				

RESULT 10
 US-08-480-229C-20
 : Sequence 20, Application US/08480229C
 : Patent No. 5874562
 :
 : GENERAL INFORMATION:
 : APPLICANT: Quettermous, Thomas
 : APPLICANT: Hogan, Bridg
 : APPLICANT: Snodgrass, H. Ralph
 : APPLICANT: Zupancic, Thomas J.
 : TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 : TITLE OF INVENTION: CELL LOCUS-1
 : NUMBER OF SEQUENCES: 29
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennile & Edmonds LLP
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: United States
 : ZIP: 10036-2711
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/480.229C
 : FILING DATE: 07-JUN-1995
 : CLASSIFICATION: 536
 :
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Polsant, Brian M.
 : REGISTRATION NUMBER: 28,462
 : REFERENCE/DOCKET NUMBER: 8907-0026-999
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (212) 790-9090
 : TELEFAX: (212) 869-8864/9741
 : TELEX: 66141 Pennile
 :
 : INFORMATION FOR SEQ ID NO: 20:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 320 amino acids
 : TYPE: amino acid
 : STRANDEDNESS:
 : TOPOLOGY: unknown
 :
 : MOLECULE TYPE: protein
 :
 : US-08-480-229C-20

[illegible]

Db 301 RVLPSVSWNRITLRLLELLGC 320

RESULT 11

US-08-659-235C-20
Sequence 20, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELEPHONE/COMMUNICATION INFORMATION:
TELEPHONE: (212) 790-8090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-659-235C-20

Query Match 35.4%; Score 948; DB 2; Length 320;

Best Local Similarity 54.7%; Pred. No. 8.5e-70; Mismatches 95; Indels 0; Gaps 0;

Db 158 CSEPLIGISGIIISMOQTASSSTRRLALGLQKWPYVARLKNKGLINAMTAENDNRKMT 217
1 CSTQLMESGALADSQISASTYVFMGLOKRWGPRLALRYGTIVAMASINADSKPMI 60
QY 218 QINLQKRMKVTGVTGAKRIGSPKRYKTIANSNDGKTWAMKVKGTNEDMVRGNID 277
Db 61 QVNLKRMKVSVMTGASRAGAEVLTKEFYAVSLDGKRFEEIDEGSGKREFLGMD 120
QY 278 NNTPVANSEFTPIKAQYVRLYPQVCRRCRCTLRMELLCGELSGCSEPLGKMSGHIDVYT 337
Db 121 NNSLKVMNEMPTLEAOYIRLYPVSCRGTLEFELLCGELHGLCLPGLKNTTIPDSQMS 180
QY 338 ASSIFETLMDFTWEPKRALDKQKRVNMTSGHNDOSQWLQVDLLVPTKYTGITITGA 397
Db 181 ASSSYTWNLRKFGYRHLGRLDNGKINAMTAOSNKEWLDVGLGTROYTGITITGA 240
QY 398 KDSGHVQFVGSYKRLAYSNDGHEHTVYQDERQKDKVFGNFDNDTHRKNVJDPPIYAH 457
Db 241 RDPGHIOVEYSKYVAHSDGVQWTVYXXEEOGSSKVFQGNLDNNSHKNIPEKPMARYV 300

QY 458 RILFWSWYGRITFLASELLGC 477
Db 301 RVLPSVSWNRITLRLLELLGC 320

RESULT 12

US-07-607-538C-3
Sequence 3, Application US/07607538C
Patent No. 5455031
GENERAL INFORMATION:
APPLICANT: Ceriani Dr., Roberto L.
APPLICANT: Peterson Dr., Jerry A.
APPLICANT: Larocca, David J.
TITLE OF INVENTION: POLYPEPTIDE WITH 46
TITLE OF INVENTION: DIFFERENTIATION ANTIGEN BINDING SPECIFICITY AND CLOTTING
TITLE OF INVENTION: FACTORS V AND VIII LIGHT-CHAIN HOMOLOGIES,
TITLE OF INVENTION: FUSION PROTEIN, POLYNUCLEOTIDE AND POLYRIBO-
TITLE OF INVENTION: NUCLEOTIDE ENCODING THE POLYPEPTIDE, ANTI-
TITLE OF INVENTION: POLYPEPTIDE ANTIBODIES, KITS AND METHODS OF
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: V. Amzel & Assoc.
STREET: 2055 No. 5455031th Broadway
CITY: Walnut Creek
STATE: California
COUNTRY: USA
ZIP: 94596
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/607,538C
FILING DATE: 01-NOV-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: CRFCC-004
TELEPHONE/COMMUNICATION INFORMATION:
TELEPHONE: (510) 943-1931
TELEFAX: (510) 943-1189
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE:
US-07-607-538C-3

Query Match 24.0%; Score 643; DB 1; Length 217;

Best Local Similarity 56.8%; Pred. No. 3.7e-45; Mismatches 59; Indels 0; Gaps 0;

Db 272 FRGNIDNTPYANSEFTPIKAQYVRLYPQVCRRCRCTLRMELLCGELSGCSEPLGKMSGH 331
12 EVGMNNKNAHVMLFTPTPEAOYVRLYPTSCHTACTLRELLGCELNCAAPLGKNNST 71
QY 332 ODYQITASSIFETLMDFTWEPKRALDKQKRVNMTSGHNDOSQWLQVDLLVPTKYTG 391
Db 72 PKQITASSSKKTKWJHLFSWNPYARLDQGNFNMAVSGYNQWLQVDLSSSEVIG 131
QY 392 IITGAKDPGHVQFVGSYKRLAYSNDGHEHTVYQDERQKDKVFGNFDNDTHRKNVJDP 451
Db 132 IITGAKDPGHVQFVGSYKRLAYSNDGHEHTVYQDERQKDKVFGNFDNDTHRKNVJDP 451

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:12:43 ; Search time 62.53 Seconds
(without alignments)
71.438 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Perfect score: 362
Sequence: 1 MKRSVAWMLVGLSLGVPQF.....FSCCEPDGFTDPNCSSVVEV 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR.67:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	131	36.2	2318	2 S45306	notch 3 protein -
2	126	34.8	2321	2 S78549	fibropellin Ia - s
3	121.5	33.6	1064	2 A40136	gene serrate prote
4	121	33.4	1408	2 S16148	transmembrane prot
5	121	33.4	2437	2 S42612	notch protein homo
6	117	32.3	2555	2 A40043	notch protein homo
7	115	31.8	2531	2 S18188	Notch-1 protein -
8	115	31.8	2531	2 A46019	fibropellin C prec
9	113.5	31.4	570	2 A48836	homeotic protein d
10	113	31.2	385	2 S53718	preadipocyte facto
11	113	31.2	385	2 A54785	Notch protein - Af
12	113	31.2	2524	2 A35844	delta-like homeoti
13	112	30.9	383	2 S53716	coagulation factor
14	112	30.9	475	1 EXCH	epidermal growth f
15	110	30.4	308	3 JC7125	hypothetical prote
16	110	30.4	3871	2 T22812	Delta-like 1 - mou
17	109	30.1	722	2 I48324	cell-fate determin
18	109	30.1	2471	2 A49128	coagulation factor
19	108	29.8	482	1 EXRT	Notch homolog prot
20	108	29.8	2352	2 T30201	Notch homolog se
21	108	29.8	2531	2 T31070	Notch homolog Motc
22	107	29.6	861	2 A48825	Notch B protein -
23	107	29.6	1203	2 A49175	notch protein - fr
24	107	29.6	2703	1 A24420	coagulation factor
25	105.5	29.1	461	1 KFHU	coagulation factor
26	104	28.7	407	1 KFB07	adhesive plaque pr
27	103.5	28.6	473	2 A56175	C-Delta-1 - chicke
28	103	28.5	728	2 I50719	homeotic protein 1
29	102.5	28.3	1429	2 S06434	

30	102.5	28.3	1964	2 T09059	notch4 - mouse
31	102.5	28.3	2809	2 T30213	C-cadherin - sea u
32	102	28.2	459	2 T00419	coagulation factor
33	102	28.2	601	2 T22025	hypothetical prote
34	102	28.2	1687	2 T30176	EGF repeat transme
35	101.5	28.0	832	2 A31246	neurogenic protein
36	101.5	28.0	880	2 S00670	neurogenic repeat
37	101	27.9	416	1 KFB0	coagulation factor
38	101	27.9	488	1 EXHU	agrin precursor -
39	101	27.9	1955	1 AGCH	hypothetical prote
40	100	27.6	1264	2 T19545	gene Delta protein
41	99.5	27.5	833	2 S19087	agrin - electric r
42	99	27.3	1328	2 T43060	jagged protein pre
43	98	27.1	1220	2 A56136	meprin A (EC 3.4.2
44	97.5	26.9	746	1 H10DMA	slic-1 protein hom
45	97	26.8	1531	2 T42218	

ALIGNMENTS

RESULT 1
S45306
notch 3 protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
R:Lardelli, M.; Dahlstrand, J.; Lendahl, U.
Mech. Dev. 46, 123-136, 1994
A:Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth fact
A:Reference number: S45306; MID:95001556
A:Accession: S45306
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2318 <L&R>
A:Cross-references: EMBL:X74760; NID:9483580; PIDN:CAA52776.1; PID:9483581
F:163-195/Domain: EGF homology <EGF>
F:163-195/Domain: EGF homology <EGF>
F:474-505/Domain: EGF homology <EGF>
F:854-885/Domain: EGF homology <EGF>
F:1839-1871/Domain: ankyrin repeat homology <AN1>
F:1839-1871/Domain: ankyrin repeat homology <AN2>
F:1872-1904/Domain: ankyrin repeat homology <AN3>
F:1906-1938/Domain: ankyrin repeat homology <AN4>
F:1939-1971/Domain: ankyrin repeat homology <AN5>
F:1972-2004/Domain: ankyrin repeat homology <AN5>

Query Match 36.2%; Score 131; DB 2; Length 2318;
Best local similarity 58.5%; Pred. No. 6.2e-06;
Matches 24; Conservative 3; Mismatches 12; Indels 2; Gaps 1;
QY 24 DICDPNCPBNGICPLGLAVGFSCECPGFTDPNCSSVVE 64
DB 852 DDCDPNCPCHGSCQDQ--VGSFSCSCLDGFGPRCARDVD 890

RESULT 2
S78549
notch3 protein - human
C:Species: Homo sapiens (man)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 20-Sep-1999
C:Accession: S78549; S71825
R:Joutel, A.; Tournier-Lasserre, E.
Submitted to the EMBL Data Library, April 1997
A:Reference number: S78549
A:Accession: S78549
A:Molecule type: mRNA
A:Residues: 1-2321 <JOU1>
A:Cross-references: EMBL:U97669; NID:92668591; PIDN:AA91371.1; PID:92668592
R:Joutel, A.; Corpechot, C.; Ducros, A.; Vahedi, K.; Chabrialat, H.; Mouton, P.; Alamow
x, M.M.; Weissbach, J.; Bach, J.F.; Bousser, M.G.; Tournier-Lasserre, E.
Nature 389, 707-710, 1996
A:Title: Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stro

A:Reference number: 571825; MUID:97032728
A:Accession: 571825
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 67-113;138-194;268-333, 'G', 335-346;536-613;716-765;1240-1279;1815-1888 <T002
A:Cross-references: EMBL:U97669
C:Genetics:
A:Gene: notch3
A:Map position: 19p13.1
C:Function:
A:Description: may be involved in pathogenesis of CADASIL, causing a type of stroke and
C:Superfamily: unassigned ankyrin repeat proteins; EGF homology
C:Keywords: tandem repeat; transmembrane protein
F:123-155/Domain: EGF homology <EGX1>
F:162-194/Domain: EGF homology <EGX1>
F:240-271/Domain: EGF homology <EGX2>
F:318-349/Domain: EGF homology <EGX>
F:473-504/Domain: EGF homology <EGX3>
F:853-884/Domain: EGF homology <EGF3>
F:928-959/Domain: EGF homology <EGX4>
F:1070-1136/Domain: laminin-type EGF-like homology <LEG>
F:1838-1870/Domain: ankyrin repeat homology <AN1>
F:1871-1903/Domain: ankyrin repeat homology <AN2>
F:1905-1937/Domain: ankyrin repeat homology <AN3>
F:1938-1970/Domain: ankyrin repeat homology <AN4>
F:1971-2003/Domain: ankyrin repeat homology <AN5>

```

Query Match Similarity      34.88; Score 126; DB 2; Length 2221;
Best Local Similarity      59.08; Pred No. 2.1e-05;
Matches 23; Conservative 2; Mismatches 12; Indels 2; Gaps 1
OY      26 CDPCPCENGAGICLPGLAVGSFSCPCDGGTTDPNCSSVYE 64
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      853 CDPCPLCNGSCDQGS--VGSFSSCTCLPGAGRCARVD 889

```

RESULT 3
A40136
fibropellin 1a - sea urchin (Strongylocentrotus purpuratus)
N:Alternate names: epidermal growth factor homolog precursor
N:Contains: alternatively spliced fibropellin 1b (EGF1)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C>Date: 13-May-1992 #sequence_revision 17-Sep-1997 #text_change 21-Jul-2000
C:Accession: A40136; B40136; C40136; A29316; A4131
R:delgadillo-Reynoso, M.G.; Rollo, D.R.; Hersh, D.A.; Reff, R.A.
J. Mol. Evol. 29, 314-327, 1989
A:Title: Structural analysis of the usgf gene in the sea urchin Strongylocentrotus purp
J:Reference number: A40136; WUID:90112459

A:Residues: 1-114
A:Cross-references: GB:117530; NID:910225; PID:9667061
A:Accession: B40136
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 181-251,329-370,'R','372-408','RA','411-441 <DE2>
A:Accession: C40136
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 'K','747-821,898-978 <DE3>
R:Hurst, D.A.; Andrews 'M.E.; Raff, R.A.
Science 237, 1487-1490, 1987
A:Title: A sea urchin gene encodes a polypeptide homologous to epidermal growth factor.
A:Reference number: A29316; KUID:87319677
A:Accession: A29316
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 'S',280-481,786-1064 <HUR>
A:Cross-references: GB:M17421; NID:9161474; PID:AAA30050.1; PID:9552260
R:Hunt, L.T.; Barker, W.C.
PASEB J. 3, 1760-1764, 1989

A:Title: Avidin-like domain in an epidermal growth factor homolog from a sea urchin.
A:Reference number: A43131. MIDB:89196806
A:Contents: annotation
C:Comment: EGF homology repeats 10-17 are spliced out in the short form (fibropellin C).Superfamily: C1r/C1s repeat homology; EGF homology
F:1-19/Domain: signal sequence status predicted <STO>
F:20-1064/Product: fibropellin 1 #status predicted <PIB>
F:23-34/Domain: EGF homology <EG31>
F:37-115/Domain: C1r/C1s repeat homology <C1R>
F:180-211/Domain: EGF homology <EG2>
F:218-249/Domain: EGF homology <EG3>
F:256-287/Domain: EGF homology <EG04>
F:294-325/Domain: EGF homology <EG05>
F:332-363/Domain: EGF homology <EG06>
F:370-401/Domain: EGF homology <EG07>
F:408-439/Domain: EGF homology <EG08>
F:446-477/Domain: EGF homology <EG09>
F:484-515/Domain: EGF homology <EG10>
F:522-553/Domain: EGF homology <EG11>
F:560-591/Domain: EGF homology <EG12>
F:598-629/Domain: EGF homology <EG13>
F:636-667/Domain: EGF homology <EG14>
F:674-705/Domain: EGF homology <EG15>
F:712-743/Domain: EGF homology <EG16>
F:750-781/Domain: EGF homology <EG17>
F:788-819/Domain: EGF homology <EG18>
F:826-857/Domain: EGF homology <EG19>
F:864-895/Domain: EGF homology <EG20>
F:902-933/Domain: EGF homology <EG21>
F:936-1064/Region: avidin-like
F:121-34, 28-43, 45-54, 62-88, 180-191, 185-200, 202-211, 218-229, 223-228, 240-249, 256-267, 261-272, 283-284, 286-287, 290-291, 293-294, 296-297, 300-301, 303-304, 306-307, 310-311, 313-314, 316-317, 319-320, 322-323, 325-326, 328-329, 331-332, 334-335, 337-338, 340-341, 343-344, 346-347, 349-350, 352-353, 355-356, 358-359, 361-362, 364-365, 367-368, 370-371, 373-374, 376-377, 379-380, 382-383, 385-386, 388-389, 391-392, 394-395, 397-398, 400-401, 403-404, 406-407, 409-410, 412-413, 415-416, 418-419, 421-422, 424-425, 427-428, 430-431, 433-434, 436-437, 439-440, 442-443, 445-446, 448-449, 451-452, 454-455, 457-458, 460-461, 463-464, 466-467, 469-470, 472-473, 475-476, 478-479, 481-482, 484-485, 487-488, 490-491, 493-494, 496-497, 499-500, 502-503, 505-506, 508-509, 511-512, 514-515, 517-518, 520-521, 523-524, 526-527, 529-530, 532-533, 535-536, 538-539, 541-542, 544-545, 547-548, 550-551, 553-554, 556-557, 559-560, 562-563, 565-566, 568-569, 571-572, 574-575, 577-578, 580-581, 583-584, 586-587, 589-590, 592-593, 595-596, 598-599, 601-602, 604-605, 607-608, 610-611, 613-614, 616-617, 619-620, 622-623, 625-626, 628-629, 631-632, 634-635, 637-638, 640-641, 643-644, 646-647, 649-650, 652-653, 655-656, 658-659, 661-662, 664-665, 667-668, 670-671, 673-674, 676-677, 679-680, 682-683, 685-686, 688-689, 691-692, 694-695, 697-698, 700-701, 703-704, 706-707, 709-710, 712-713, 715-716, 718-719, 721-722, 724-725, 727-728, 730-731, 733-734, 736-737, 739-740, 742-743, 745-746, 748-749, 751-752, 754-755, 757-758, 760-761, 763-764, 766-767, 769-770, 772-773, 775-776, 778-779, 781-782, 784-785, 787-788, 790-791, 793-794, 796-797, 799-800, 802-803, 805-806, 808-809, 811-812, 814-815, 817-818, 820-821, 823-824, 826-827, 829-830, 832-833, 835-836, 838-839, 841-842, 844-845, 847-848, 850-851, 853-854, 856-857, 859-860, 862-863, 865-866, 868-869, 871-872, 874-875, 877-878, 880-881, 883-884, 886-887, 889-890, 892-893, 895-896, 898-899, 901-902, 904-905, 907-908, 910-911, 913-914, 916-917, 919-920, 922-923, 925-926, 928-929, 931-932, 934-935, 937-938, 940-941, 943-944, 946-947, 949-950, 952-953, 955-956, 958-959, 961-962, 964-965, 967-968, 970-971, 973-974, 976-977, 979-980, 982-983, 985-986, 988-989, 991-992, 994-995, 997-998, 1000-1001, 1003-1004, 1006-1007, 1009-1010, 1012-1013, 1015-1016, 1018-1019, 1021-1022, 1024-1025, 1027-1028, 1030-1031, 1033-1034, 1036-1037, 1039-1040, 1042-1043, 1045-1046, 1048-1049, 1051-1052, 1054-1055, 1057-1058, 1060-1061, 1063-1064, 1066-1067, 1069-1070, 1072-1073, 1075-1076, 1078-1079, 1081-1082, 1084-1085, 1087-1088, 1090-1091, 1093-1094, 1096-1097, 1099-1100, 1102-1103, 1105-1106, 1108-1109, 1111-1112, 1114-1115, 1117-1118, 1120-1121, 1123-1124, 1126-1127, 1129-1130, 1132-1133, 1135-1136, 1138-1139, 1141-1142, 1144-1145, 1147-1148, 1150-1151, 1153-1154, 1156-1157, 1159-1160, 1162-1163, 1165-1166, 1168-1169, 1171-1172, 1174-1175, 1177-1178, 1180-1181, 1183-1184, 1186-1187, 1189-1190, 1192-1193, 1195-1196, 1198-1199, 1201-1202, 1204-1205, 1207-1208, 1210-1211, 1213-1214, 1216-1217, 1219-1220, 1222-1223, 1225-1226, 1228-1229, 1231-1232, 1234-1235, 1237-1238, 1240-1241, 1243-1244, 1246-1247, 1249-1250, 1252-1253, 1255-1256, 1258-1259, 1261-1262, 1264-1265, 1267-1268, 1270-1271, 1273-1274, 1276-1277, 1279-1280, 1282-1283, 1285-1286, 1288-1289, 1291-1292, 1294-1295, 1297-1298, 13

```

Query MatchSimilarity    33.6%; Score 121.5; DB 2: length 1064;
Best local similarity    46.38% Pred. No. 3.2e-05;
Matches    25; Conservative    7; Mismatches    17; Indels    5; Gaps    3;

OY      8 WLLAVGLSLGV--PQFGKGDICDPNCEKGSGIGCLPLGVLNGVSFSSECPDGFPTTPNC 59
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      4 WLLAVLTLVLAIVTYGGGE-CDSDFPCENGSTCGES--EGSYIQCCPMGIQDQC 54

RESULT      4
S16148
gene serrate protein precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 31-Dec-1991 #sequence_revision 02-Aug-1994 #ext_change 17-Nov-2000
C:Accession: S16148; S16878; A36666
R:Thomas, U.; Speicher, S.A.; Knust, E.
Development 111, 749-761, 1991
A>Title: The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a
A:Reference number: S16148; MUID:91347903
A:Accession: S16148
A:Molecule type: mRNA
A:Residues: 1-1408 <THOI>
A:Cross-references: EMBL:X56811
R:Thomas, U.
submitted to the EMBL Data Library, November 1990
A:Reference number: S16878
A:Accession: S16878
A:Molecule type: mRNA
A:Residues: 1-1351,'T','I351-1408 <THO2>
A:Cross-references: EMBL:X56811; NID:89563; PID:98564
R:Fleming, R.J.; Scottsdale, T.N.; Diederich, R.J.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2188-2201, 1990
A>Title: The gene Serrate encodes a putative EGF-like transmembrane protein essential
A:Reference number: A36666; MUID:91099666
A:Accession: A36666
A>Status: preliminary
A:Molecule type: mRNA

```

A:Residues: 1-15,20-26,'A',28-1408 <FILE>
 A:Cross-References: GB:M35759; NID:9158605; PID:9158606
 C:Genetics:
 A:Gene: FlyBase:Ser
 A:Cross-References: FlyBase:FBgn0004197
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: glycoprotein; transmembrane protein
 F:1-84/Domain: signal sequence <status predicted <SIG>
 F:85-1408/Product: gene serrate protein <status predicted <NNT>
 F:85-1408/Domain: extracellular <status predicted <EXT>
 F:283-316/Domain: EGF homology <EG01>
 F:319-348/Domain: EGF homology <EG02>
 F:355-388/Domain: EGF homology <EG03>
 F:395-488/Domain: EGF homology <status atypical <EG04>
 F:495-526/Domain: EGF homology <EG05>
 F:533-608/Domain: EGF homology <status atypical <EG06>
 F:615-645/Domain: EGF homology <EG07>
 F:652-683/Domain: EGF homology <EG08>
 F:690-720/Domain: EGF homology <EG09>
 F:727-796/Domain: EGF homology <status atypical <EG10>
 F:803-834/Domain: EGF homology <EG11>
 F:841-876/Domain: EGF homology <EG12>
 F:883-914/Domain: EGF homology <EG13>
 F:921-952/Domain: EGF homology <EG14>
 F:997-1060/Region: cysteine-rich
 F:1222-1246/Domain: transmembrane <status predicted <TM1>
 F:1247-1408/Domain: intracellular <status predicted <INT>
 F:152,196,247,331,412,452,558,739,965,977,1004,1030,1150/Binding site: carbohydrate (asn

Query Match 33.4%; Score 121; DB 2; Length 1408;
 Best Local Similarity 60.0%; Pred. No. 4.5e-05;
 Matches 21; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

26 CDPNPCENGICLPGLAVGFSCECPDPTDPCNS 60
 DB 803 CSPNRCNGICLDG--DGDFTCECMSCWTKRCS 835

RESULT 5
 S42612

transmembrane protein precursor - zebra fish
 C:Species: Brachydanio rerio (zebra fish)
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
 C:Accession: S42612
 R:Bierkamp, C.; Campos-Ortega, J.A.
 Mech. Dev. 43, 87-100, 1993
 A:Title: A zebrafish homologue of the Drosophila neurogenic gene Notch and its pattern
 A:Reference number: S42612; MUID:94128602
 A:Accession: S42612
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2437 <BIE>
 A:Cross-References: EMBL:X69088; NID:9433866; PID:CAA4881.1; PID:9433867
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
 F:755-786/Domain: EGF homology <EG1>
 F:1023-1054/Domain: EGF homology <EGF>
 F:1185-1216/Domain: EGF homology <EGF2>
 F:1917-1947/Domain: ankyrin repeat homology <AN1>
 F:1948-1980/Domain: ankyrin repeat homology <AN2>
 F:1982-2014/Domain: ankyrin repeat homology <AN3>
 F:2015-2047/Domain: ankyrin repeat homology <AN4>
 F:2048-2080/Domain: ankyrin repeat homology <AN5>

Query Match 33.4%; Score 121; DB 2; Length 2437;
 Best Local Similarity 57.9%; Pred. No. 7.3e-05;
 Matches 22; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

24 DICDPNCPENGICLPGLAVGFSCECPDPTDPCNS 61
 DB 678 DDCALNPFCHNGICTDGG--VNSFTCLCPDGFDPATCLS 713

RESULT 6
 A40043
 notch protein homolog TAN-1 precursor - human
 C:Species: Homo sapiens (man)
 C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 13-Aug-1999
 C:Accession: A40043
 R:Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; SKI
 Cell 66, 649-661, 1991
 A:Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromoso
 A:Reference number: A40043; MUID:91347367
 A:Accession: A40043
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-2555 <EUL>
 A:Cross-References: GB:M73980
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
 F:261-292/Domain: EGF homology <EGX1>
 F:494-525/Domain: EGF homology <EGF1>
 F:987-1018/Domain: EGF homology <EGX2>
 F:1149-1180/Domain: EGF homology <EGF>
 F:1187-1218/Domain: EGF homology <EGF3>
 F:1233-1264/Domain: EGF homology <EGX3>
 F:1927-1959/Domain: ankyrin repeat homology <AN1>
 F:1960-1992/Domain: ankyrin repeat homology <AN2>
 F:1994-2026/Domain: ankyrin repeat homology <AN3>
 F:2027-2059/Domain: ankyrin repeat homology <AN4>
 F:2060-2092/Domain: ankyrin repeat homology <AN5>

Query Match 32.3%; Score 117; DB 2; Length 2555;
 Best Local Similarity 50.0%; Pred. No. 0.0002;
 Matches 20; Conservative 4; Mismatches 14; Indels 2; Gaps 1;

24 DICDPNCPENGICLPGLAVGFSCECPDPTDPCNS 63
 DB 679 DECAGNPFCHNGICTEDG--INGFTCRCEGYHDPICLSEV 716

RESULT 7
 S18188

notch protein homolog - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
 C:Accession: S18188
 R:Welmaster, G.; Roberts, V.J.; Lemke, G.
 Development 113, 199-205, 1991
 A:Title: A homolog of Drosophila Notch expressed during mammalian development.
 A:Reference number: S18188; MUID:92111383
 A:Accession: S18188
 A:Molecule type: mRNA
 A:Residues: 1-2531 <WEI>
 A:Cross-References: EMBL:X57405; NID:957634; PID:957635
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
 F:987-1018/Domain: EGF homology <EGF1>
 F:1023-1056/Domain: EGF homology <EGF>
 F:1233-1264/Domain: EGF homology <EGF2>
 F:1917-1949/Domain: ankyrin repeat homology <AN1>
 F:1950-1982/Domain: ankyrin repeat homology <AN2>
 F:1984-2016/Domain: ankyrin repeat homology <AN3>
 F:2017-2049/Domain: ankyrin repeat homology <AN4>
 F:2050-2082/Domain: ankyrin repeat homology <AN5>

Query Match 31.8%; Score 115; DB 2; Length 2531;
 Best Local Similarity 50.0%; Pred. No. 0.00032;
 Matches 20; Conservative 5; Mismatches 13; Indels 2; Gaps 1;

24 DICDPNCPENGICLPGLAVGFSCECPDPTDPCNS 63
 DB 680 DECAGSPCHNGICTEDGTA--GFTCRCEGYHDPICLSEV 717

```

RESULT      8
A46019
Notch-1 protein - mouse
N:Alternate names: notch protein
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C:Accession: A46019; S25144
Ridel Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Grid
Genomics 15, 259-264, 1993
A:Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of
A:Reference number: A46019; MUID:93194170
A:Accession: A46019
A:Status: not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-2531 <DEL>
A:Cross-references: GB:211886; GB:S47228; NID:q288502; PIDN:CAAT7941.1; PID:q288503
A>Note: sequence extracted from NCBI backbone (NCBI:P127318)
R.Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.;
submitted to the EMBL Data Library, April 1992
A:Description: Expression pattern of Notch, a mouse homolog of Drosophila Notch, suggest
A:Reference number: S25144
A:Accession: S25144
A:Molecule type: mRNA
A:Residues: 1551-2108, 'Q', 2110-2114, 'ALP', 2118-2170 <FRA>
A:Cross-references: EMBL:Z11886
C:Genetics:
A:Gene: notch-1
A:Map position: 2
A>Note: proximal region of chromosome 2
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:106-138/Domain: EGF homology <EGF1>
F:144-175/Domain: EGF homology <EGO1>
F:222-254/Domain: EGF homology <EGF2>
F:261-292/Domain: EGF homology <EGO2>
F:339-370/Domain: EGF homology <EGO3>
F:416-449/Domain: EGF homology <EGO3>
F:456-487/Domain: EGF homology <EGO4>
F:494-525/Domain: EGF homology <EGO5>
F:532-563/Domain: EGF homology <EGO6>
F:607-638/Domain: EGF homology <EGO7>
F:682-713/Domain: EGF homology <EGO8>
F:757-788/Domain: EGF homology <EGO9>
F:795-826/Domain: EGF homology <EGO9>
F:873-904/Domain: EGF homology <EG11>
F:911-942/Domain: EGF homology <EG11>
F:949-980/Domain: EGF homology <EG12>
F:987-1018/Domain: EGF homology <EG13>
F:1025-1056/Domain: EGF homology <EG14>
F:1063-1094/Domain: EGF homology <EG16>
F:1149-1180/Domain: EGF homology <EG17>
F:1187-1218/Domain: EGF homology <EG18>
F:1233-1264/Domain: EGF homology <EGF4>
F:1352-1383/Domain: EGF homology <EG19>
F:1391-1425/Domain: EGF homology <EGF>
F:1917-1948/Domain: ankyrin repeat homology <AN1>
F:1949-1981/Domain: ankyrin repeat homology <AN2>
F:1983-2015/Domain: ankyrin repeat homology <AN3>
F:2016-2048/Domain: ankyrin repeat homology <AN4>
F:2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match      31.88; Score 115; DB 2; Length 2531;
Best Local Similarity 50.0%; Pred. No. 0.00032;
Matches 20; Conservative 5; Mismatches 13; Indels 2; Gaps 1;

OY      24 DICDPNCGGICLPGLAVGSFSCPCPDGFTDPNCSVY 63
DB      688 DECAAGSPCHNGICEDGIA--GFTCRCPREYHDPFCLSEY 717

RESULT      9
A48836
fibropellin C precursor - sea urchin (Strongylocentrotus purpuratus)

```

```

N:Alternate names: EGF repeat-containing protein; epidermal growth factor-related pro
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C:Accession: A48836
R:Bisgrove, B.W.; Rafi, R.A
Dev. Biol. 157, 526-538, 1993
A:Title: The SpEGF III gene encodes a member of the fibropellins: EGF repeat-containi
A:Reference number: A48836; MUID:93273088
A:Accession: A48836
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-570 <BIS>
A:Cross-references: GB:L07045; NID:q310659; PID:q310660
A>Note: sequence extracted from NCBI backbone (NCBI:N132724, NCBI:P132725)
C:Superfamily: C1r/C1s repeat homology; EGF homology
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-570/Product: fibropellin C #status predicted <FIB>
F:19-54/Domain: EGF homology <EG1>
F:57-175/Domain: C1r/C1s repeat homology <C1R>
F:176-211/Domain: EGF homology <EG2>
F:214-249/Domain: EGF homology <EG3>
F:252-287/Domain: EGF homology <EG4>
F:290-325/Domain: EGF homology <EG5>
F:328-363/Domain: EGF homology <EG6>
F:366-401/Domain: EGF homology <EG7>
F:404-439/Domain: EGF homology <EG8>
F:442-570/Region: avidin-like
F:23-34,28-43,45-54,62-88,180-191,185-200,202-211,218-229,223-238,240-249,256-267,261
tide bonds: #status predicted

Query Match      31.4%; Score 113.5; DB 2; Length 570;
Best Local Similarity 42.4%; Pred. No. 0.00013;
Matches 25; Conservative 9; Mismatches 20; Indels 5; Gaps 3;

OY      1 MKRSVAWMLVGLSIGVPGFKGICDPNCGGICLPGLAVGSFSCPCPDGFTDPNCSVY 59
DB      1 MKVSLVAVLL--LSTVATVYGGG--CGSNPCNCSVCRDS--EGTVICRCOMGYDQNC 54

RESULT      10
S53718
homeotic protein dlx - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 11-Jan-2000
C:Accession: S53718
R:Lee, Y.L.; Helman, L.; Hoffman, T.; Laborda, J.
Biochim. Biophys. Acta 1261, 223-232, 1995
A:Title: dlx, pg2 and Pref-1 mRNAs encode similar proteins belonging to the EGF-like
A:Reference number: S53716; MUID:95226449
A:Accession: S53718
A:Molecule type: mRNA
A:Residues: 1-385 <LEB>
A:Cross-references: EMBL:U15980; NID:q562107; PIDN:AB60495.1; PID:q562108
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: transmembrane protein
F:54-85/Domain: EGF homology <EGF1>
F:92-124/Domain: EGF homology <EGF>
F:131-167/Domain: EGF homology <EGF3>
F:214-246/Domain: EGF homology <EGX1>
F:303-332/Domain: transmembrane #status predicted <TM1>

Query Match      31.2%; Score 113; DB 2; Length 385;
Best Local Similarity 52.6%; Pred. No. 0.0001;
Matches 20; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

OY      26 CDPNCGGICLPGLAVGSFSCPCPDGFTDPNCSVY 63
DB      176 CTPNCPENDGVCTD--IGDFRCRCPAGFVDTCSRPV 211

RESULT      11

```

A54785
 preadipocyte factor 1 precursor, long form - mouse
 N:Alternate names: delta-like dlk homeotic protein; pref-1
 C:Species: Mus musculus (house mouse)
 C:Date: 12-Apr-1995 #sequence,revision 12-Oct-1995 #text,change 11-Jan-2000
 C:Accession: A54785; A45484; A40746; S21585
 R:Smas, C.M.; Green, D.; Sul, H.S.
 Biochemistry 33, 9257-9265, 1994
 A:Title: Structural characterization and alternate splicing of the gene encoding the preadipocyte factor 1 precursor, long form - mouse
 A:Reference number: A54785; MUID:94325252
 A:Accession: A54785
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-365 <SMA>
 A:Cross-references: GB:S71340
 R:Laborda, J.; Sauvillie, E.A.; Hoffman, T.; Notario, V.
 J. Biol. Chem. 268, 3817-3820, 1993
 A:Title: dlk, a putative mammalian homeotic gene differentially expressed in small cell
 A:Reference number: A45484; MUID:93179372
 A:Accession: A45484
 A:Molecule type: mRNA
 A:Residues: 1-78; 'G', 80-343; 'T', 346-385 <LAB>
 A:Cross-references: EMBL:212171; NID:950716; PIDN:CA478162.1; PID:950717
 A:Note: sequence extracted from NCBI backbone (NCBI:125734)
 A:Note: species designations for this sequence report and for B45484 originally were the R:Smas, C.M.; Sul, H.S.
 Cell 73, 725-734, 1993
 A:Title: Pref-1, a protein containing EGF-like repeats, inhibits adipocyte differentiation
 A:Reference number: A40746; MUID:93272313
 A:Accession: A40746
 A:Molecule type: mRNA
 A:Residues: 1-78; 'G', 80-249; 'P', 251-319; 'C', 320-385 <LAB>
 A:Cross-references: GB:112721; NID:9309092; PIDN:AAA37175.1; PID:9309093
 A:Note: experimental source: j37-11 preadipocytes
 A:Note: this sequence appears to have been corrected in reference A45484
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; tandem repeat; transmembrane protein
 F:1-385/Product: preadipocyte factor 1 precursor splice form A #status predicted <RAN>
 F:1-230,283-385/Product: preadipocyte factor 1 precursor splice form B #status predicted
 F:1-230,304-385/Product: preadipocyte factor 1 precursor splice form C #status predicted
 F:1-230,306-385/Product: preadipocyte factor 1 precursor splice form D #status predicted
 F:1-210,306-385/Product: preadipocyte factor 1 precursor splice form D #status predicted
 F:92-124/Domain: EGF homology <EGF1>
 F:214-246/Domain: EGF homology <EGF1>

Query Match 31.2% Score 113; DB 2; Length 385;
 Best Local Similarity 52.6% Pred. No. 0.0001;
 Matches 20; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

OY 26 CDPNCPENGICLPGIAGVSCCECPDGFDPNCSSV 63
 DB 176 CTPNCPENGICVTD--IGDPRCNCAGFVDRKCSRPV 211

RESULT 12
 A5844
 Notch protein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 12-Oct-1990 #sequence,revision 12-Oct-1990 #text,change 13-Aug-1999
 C:Accession: A35844
 R:Coffman, C.; Harris, W.; Kintner, C.
 Science 249, 1438-1441, 1990
 A:Title: Notch, the Xenopus homolog of Drosophila notch.
 A:Reference number: A35844; MUID:90385285
 A:Accession: A35844
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-2524 <COP>
 C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
 C:Keywords: transmembrane protein

F:146-177/Domain: EGF homology <EGX1>
 F:184-215/Domain: EGF homology <EGF1>
 F:222-254/Domain: EGF homology <EGF>
 F:456-487/Domain: EGF homology <EGX2>
 F:757-788/Domain: EGF homology <EGF3>
 F:1025-1056/Domain: EGF homology <EGX3>
 F:1924-1956/Domain: ankyrin repeat homology <AN1>
 F:1957-1989/Domain: ankyrin repeat homology <AN2>
 F:1991-2023/Domain: ankyrin repeat homology <AN3>
 F:2024-2056/Domain: ankyrin repeat homology <AN4>
 F:2057-2089/Domain: ankyrin repeat homology <AN5>

Query Match 31.2% Score 113; DB 2; Length 2524;
 Best Local Similarity 52.8% Pred. No. 0.00052;
 Matches 19; Conservative 1; Mismatches 14; Indels 2; Gaps 1;

OY 24 DCDPNCENGICLPGIAGVSCCECPDGFDPNC 59
 DB 909 DCDPNCENGICSDG--IMFPCNCPAGFRGPKC 942

RESULT 13
 S53716
 delta-like homeotic protein dlk, long splice form precursor - human
 N:Alternate names: fetal antigen 1 (FA1)
 N:Contains: delta-like homeotic protein dlk, short splice form
 C:Species: Homo sapiens (man)
 C:Date: 06-Dec-1996 #sequence,revision 06-Dec-1996 #text,change 05-May-2000
 C:Accession: S53716; S53717; S71068; B45484; S48713; A44549; S31973; S31974
 R:Lee, Y.L.; Helman, L.; Hoffman, T.; Laborda, J.
 Biochim. Biophys. Acta 1261, 223-232, 1995
 A:Title: dlk, p62 and Pref-1 mRNAs encode similar proteins belonging to the EGF-like
 A:Reference number: S53716; MUID:95226449
 A:Accession: S53716
 A:Molecule type: mRNA
 A:Residues: 1-383 <LEP1>
 A:Cross-references: EMBL:U15979; NID:9562105; PIDN:AAA75364.1; PID:9562106
 A:Note: the authors translated the codon CAG for residue 46 as His and CCG for residu
 A:Note: the sequence in Genbank entry HS015979 has a 3 base deletion mutation in the
 A:Accession: S53717
 A:Molecule type: mRNA
 A:Residues: 1-228,302-383 <LEP2>
 A:Cross-references: EMBL:U15981; NID:9562109; PIDN:AAA75365.1; PID:9562110
 R:Laborda, J.
 submitted to the EMBL Data Library, October 1994
 A:Reference number: S71068
 A:Accession: S71068
 A:Molecule type: mRNA
 A:Residues: 1-344,346-383 <LAB1>
 A:Cross-references: EMBL:U15979; NID:9562105; PIDN:AAA75364.1; PID:9562106
 R:Laborda, J.; Sauvillie, E.A.; Hoffman, T.; Notario, V.
 J. Biol. Chem. 268, 3817-3820, 1993
 A:Title: dlk, a putative mammalian homeotic gene differentially expressed in small ce
 A:Reference number: A45484; MUID:93179372
 A:Accession: B45484
 A:Molecule type: mRNA
 A:Residues: 1-45; 'HV', 48-383 <LAB2>
 A:Cross-references: EMBL:212172; NID:938478; PIDN:CA478163.1; PID:938479
 A:Note: sequence extracted from NCBI backbone (NCBI:125735)
 R:Jensen, C.H.; Krogh, T.; Hojrup, P.; Clausen, P.P.; Skjold, K.; Larsson, L.I.; En
 Eur. J. Biochem. 225, 83-92, 1994
 A:Title: Protein structure of fetal antigen 1 (FA1). A novel circulating human epider
 A:Reference number: S48713; MUID:95010145
 A:Accession: S48713
 A:Molecule type: protein
 A:Residues: 24-107; 'D', 109-282 <JEN>
 R:Hojrup, P.; Jensen, C.H.; Skjold, K.; Teisner, B.
 Protein Sci. 2(Suppl.1), 259S, 1993
 A:Title: Primary structure of human fetal antigen 1 (FA1), a putative homeotic glycop
 A:Reference number: A44549
 A:Accession: A44549

Amptgen & Co

1

May 1951

95452

length: 2000000000

Using: Minimum Match 0%

Listing first 45 summaries

SWISSFIDOL-33 :
BUCHSSE :

and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description
1	131	36.2	2318	1	NTC3_MOUSE	Q61982 mus musculus
2	121.5	33.6	1064	1	FBP1_STRPU	P10079 strongylococ
3	121	33.4	1408	1	SERR_DROME	P18168 drosophila
4	121	33.4	2437	1	NTC3_BHARE	P46350 brachydanio
5	117	32.3	2444	1	NTC1_HUMAN	P46531 homo sapiens
6	117	31.8	2531	1	NTC1_MOUSE	Q01705 mus musculus
7	115	31.0	1231	1	FBP3_STRPU	P19019 rat
8	113.5	31.4	5570	1	DL1_MOUSE	Q09163 mus musculus
9	113	31.2	385	1	DL1_MOUSE	Q00548 homo sapiens
10	113	31.2	7233	1	NTC3_HUMAN	P21783 xenopus lae
11	113	31.2	2524	1	NTC3_XENLA	P80370 homo sapiens
12	112	30.9	383	1	DLK_HUMAN	P80370 homo sapiens
13	112	30.9	475	1	FA10_CHICK	P25155 gallus galli
14	109	30.1	722	1	DL1_MOUSE	P61483 mus musculus
15	108	29.8	714	1	DL1_RAT	P97677 rattus norv
16	106	29.3	569	1	DL1_RAT	P088671 rattus norv
17	105.5	29.1	461	1	FA9_HUMAN	P00740 homo sapiens
18	104	28.7	457	1	FA7_BOVIN	P22457 bos taurus
19	103.5	28.6	471	1	FP2_MYGA	Q25464 mytilus gal
20	103	28.5	430	1	FA10_RABIT	P07655 mus musculus
21	103	28.5	1480	1	NTC3_MOUSE	P07655 mus musculus
22	103	28.5	2703	1	NTC3_MOUSE	P14585 caecopoda
23	102.5	28.3	1459	1	IL12_CAROL	P16284 mus musculus
24	102	28.3	459	1	FA9_BOVIN	P00741 bos taurus
25	101	27.9	416	1	FA9_BOVIN	P00741 bos taurus
26	101	27.9	488	1	FA10_HUMAN	P00742 homo sapiens
27	101	27.9	1955	1	AGRI_CHICK	P1696 gallus galli
28	99.5	27.5	833	1	DL_DROME	P10041 discophya c
29	99	27.3	1338	1	AGRI_DISOM	Q50404 discophya c
30	97.5	26.9	746	1	MEPA_HUMAN	Q16819 homo sapiens
31	96	26.5	452	1	FA10_BOVIN	P00743 bos taurus
32	96	26.5	592	1	DL3_MOUSE	P088516 mus musculus
33	96	26.5	1959	1	AGRI_RAT	P25043 rattus norv

[illegible]

ALIGNMENTS

```

RESULT 1
NTC3_MOUSE STANDARD: PRT: 2318 AA.
ID NTC3_MOUSE
OC 061862:1987 (Rel. 35, Created)
PT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH 3 PROTEIN.
GN NOTCH3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TextID:100900;
DR STRAIN-ICR X SWISS WEBSTER;
RC SEQUENCE FROM N.A.
RA Larelli M., Balstrand J., Lendahl U.;
MEDLINE:95001556; PubMed:7918097;
RT "The novel Notch homologue mouse Notch 3 lacks specific epidermal
RT growth factor-repeats and is expressed in proliferating
RT neuroepithelium".
RL Mech. Dev. 46:123-136(1994).
CC 1- FUNCTION: NOTCH 3 IS A CELL SURFACE RECEPTOR INVOLVED IN THE
CC DEVELOPING CNS AND PROBABLY OTHER REGIONS OF THE EMBRYO.
CC 2- TISSUE SPECIFICITY: PROLIFERATING NEUROEPITHELIUM.
CC 3- DEVELOPMENTAL STAGE: CNS DEVELOPMENT.
CC 4- SIMILARITY: CONTAINS 34 EGF-LIKE DOMAINS.
CC 5- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC 6- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 7- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 8- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 9- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 10- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 11- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 12- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 13- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 14- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 15- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 16- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 17- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 18- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 19- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 20- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 21- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 22- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 23- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 24- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 25- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 26- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 27- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 28- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 29- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 30- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 31- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 32- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 33- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 34- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 35- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 36- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 37- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 38- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 39- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 40- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 41- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 42- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 43- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 44- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 45- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 46- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 47- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 48- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 49- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 50- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 51- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 52- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 53- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 54- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 55- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 56- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 57- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 58- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 59- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 60- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 61- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 62- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 63- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 64- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 65- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 66- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 67- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 68- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 69- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 70- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 71- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 72- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 73- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 74- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 75- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 76- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 77- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 78- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 79- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 80- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 81- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 82- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 83- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 84- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 85- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 86- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 87- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 88- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 89- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 90- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 91- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 92- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 93- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 94- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 95- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 96- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 97- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 98- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 99- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 100- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 101- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 102- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 103- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 104- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 105- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 106- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 107- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 108- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 109- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 110- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 111- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 112- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 113- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 114- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 115- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 116- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 117- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 118- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 119- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 120- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 121- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 122- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 123- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 124- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 125- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 126- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 127- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 128- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 129- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 130- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 131- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 132- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 133- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 134- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 135- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 136- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 137- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 138- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 139- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 140- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 141- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 142- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 143- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 144- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 145- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 146- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 147- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 148- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 149- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 150- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 151- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 152- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 153- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 154- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 155- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 156- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 157- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 158- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 159- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 160- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 161- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 162- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 163- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 164- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 165- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 166- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 167- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 168- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 169- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 170- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 171- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 172- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 173- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 174- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 175- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC 176
```



```

      d 1; length 2318;
      2e-06;
      12; indels 2; gaps 1;
      64
  
```

us-09-237-981-14_copy_33_97.rsp

Page 2

```

Med May 23 06:25:14
      4 1; Length 2318;
      2e-06;
      12; Indels      2; Gaps
      12;
      36.2%; SCDFPCSSVVE 64
      58.5%; GCGAGGRCARDVD 890
Query Match Similarity
Best Local 24; Conservative
Matches
      24 DCCDPCNCGICG
      400 PRT; 1064 AA.
      852 DCCDPCNCGICG (created)
      400 (sequence update)
      400 (last annotation update)
Db CURSOR (EPIDERMAL GROWTH FACTOR-RELATED PROTEIN 1)
RESULT
F1: Syntrophus purpuratus (Purple sea urchin)
ID: Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Acta; Euechinozoa; Echinoida; Strongylocentrotidae;
Locentrotus.
Accession: J0668;

```

CC CC JOURNAL FROM N.A. PubMed-25414373;
CC CC MEDLINE-873196677; PubMed-4498216;
CC CC Delgado-Alila-Reynoso M.G., Rollo D. R., Hursh D.A., Raff R.A. :
CC CC Structural analysis of the uBFG gene in the sea urchin
CC CC strongylocentrotus purpuratus reveals more similarity to vertebrate
CC CC than to invertebrate genes with EGF-like repeats.";
CC CC J. Mol. Evol. 29:314-337(1989).
CC CC
CC CC SOURCE OF 275-476 AND 781-1064 FROM N.A.
CC CC MEDLINE-873196677; PubMed-4498216;
CC CC Hursh D.A., Andrews M.E., Raff R.A. :
CC CC "A sea urchin gene encodes a polypeptide homologous to epidermal
CC CC growth factor.";
CC CC Science 237:1487-1490(1987).
CC CC
CC CC AVIDIN-LIKE DOMAIN.
CC CC MEDLINE-89196806; PubMed-2784773;
CC CC Hunt L.T., Barker W.C. :
CC CC "Avidin-like domain in an epidermal growth factor homolog from a sea
CC CC urchin".
CC CC PNASB J. 3:1760-1764(1989).
CC CC
CC CC CHARACTERIZATION.
CC CC MEDLINE-91285254; PubMed-2060714;
CC CC Bisgrove B.W., Andrews M.E., Raff R.A. :
CC CC "Fibropellins, products of an EGF repeat-containing gene, form a
CC CC unique extracellular matrix structure that surrounds the sea urchin
CC CC embryo".
CC CC J. Biol. Chem. 264:14679-14687(1989).
CC CC
CC CC -1- FUNCTION: FORM THE APICAL LAMINA, A COMPONENT OF THE EXTRACELLULAR
CC CC MATRIX.
CC CC
CC CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR, IN VESICLES IN THE CYTOPLASM
CC CC OF UNFERTILIZED EGGS, THEN TO THE BASE OF THE HYALIN LAYER
CC CC THROUGHOUT DEVELOPMENT AND FINALLY IN THE APICAL LAMINA IN LATE
CC CC DEVELOPMENTAL STAGES.
CC CC
CC CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: IA (SHOWN HERE) AND IB: LATE
CC CC PRODUCED BY ALTERNATIVE SPLICING. THE SMALL FORM (IB) LACKS 8 EGF
CC CC REPEATS.
CC CC
CC CC -1- DEVELOPMENTAL STAGE: MODERATE LEVELS IN UNFERTILIZED EGGS AND
CC CC DURING EARLY CLEAVAGE, THEN RAPIDLY INCREASES IN ABUNDANCE BETWEEN
CC CC MATERNAL AND ZYGOTIC STAGES, THROUGHOUT DEVELOPMENT, TO MAXIMAL LEVELS
CC CC MATERNAL AND ZYGOTIC STAGES. EXPRESSED BOTH MATERNALLY
CC CC AND ZYGOTICALLY.
CC CC
CC CC -1- SIMILARITY: CONTAINS 21 EGF-LIKE DOMAINS.
CC CC
CC CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
CC CC
CC CC -1- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR
CC CC TO AVIDIN/STREPTAVIDIN.

[illegible]

[illegible]

```

DN SERATE PROTEIN PRECURSOR (BEADED PROTEIN) .
OS SER OR BD.
OS Drosophila melanogaster ('fruit fly').
OS Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OS Euryptera; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha,
OS Eurypteroidea; Eurypteroidea; Drosophila.
OX NCBI_TaxID=7227.
RN RP SEQUENCE FROM N.A.
RC STRAIN-OREGON-R;
RC MEDLINE-91347903; PubMed-1840519;
RC Thomas U.; Spelcher S.A.; Knust E.;
RC "The gene encoding a single-chain EGF-like transmembrane protein with a complex expression pattern in embryos and wing discs".
RL Development 111:749-761(1991).
RN RP [2]
RN RP SEQUENCE FROM N.A.
RC MEDLINE-91099666; PubMed-2125287;
RC Fleming R.V.; Scottgale T.N.; Diegelrich R.J.; Artavanis-Tsakonas S.;
RC "The gene encoding a putative EGF-like transmembrane protein essential for proper ectodermal development in Drosophila melanogaster".
RL Genes Dev. 4:2188-2201(1990).
CC -1 FUNCTION IS ESSENTIAL FOR PROPER ECTODERMAL DEVELOPMENT. SERRATE MAY REPRESENT AN ELEMENT IN A NETWORK OF INTERACTING MOLECULES OPERATING AT THE CELL SURFACE DURING THE DIFFERENTIATION OF CERVICAL TISSUES.
CC -1 GROWTH PROMOTION. TYPE I MEMBRANE PROTEIN.
CC -1 TISSUE SPECIFICITY: APPEARS TO BE RESTRICTED EXCLUSIVELY TO CELLS OF ECTODERMAL ORIGIN.
CC -1 MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.
CC -1 SERRATE AND NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN LEVEL. IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY COMPETE FOR BINDING WITH THE NOTCH PROTEIN.
CC -1 SIMILARITY: BELONGS TO THE DELTA/SERATE/JAGGED FAMILY.
CC -1 SIMILARITY: BELONGS TO THE DELTA/SERATE/JAGGED FAMILY.
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute and the EMBL Outstation at the European Bioinformatics Institute and its content is in no way modified by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-stb.ch).
CC CC PIR: S16811; AAA0338.1; -.
CC EMBL: M35755; AAA2938.1; -.
CC DR PIR: A36666; A36666.1; -.
CC DR PIR: S16876; S16876.
CC HSSP: P00743; IWE.
CC FLYBase: Fgn0004197; Ser.
CC FLYBase: Fgn000152; -.
CC InterPro: IPR000152; -.
CC InterPro: IPR000563; -.
CC InterPro: IPR001774; -.
CC InterPro: IPR001774; -.
CC InterPro: IPR001881; -.
CC Pfam: PF01414; DSL.1.
CC Pfam: PF00008; EGF_11.
CC PRINTS: PR00010; EGFLDXXD.
CC PROSITE: PS00010; ASX_HYDROXYL_7.
CC PROSITE: PS00022; EGF_1_34.
CC PROSITE: PS00022; EGF_1_34.
CC PROSITE: PS01187; EGF_Cn_5.
CC Differentiation: Repeat: EGF-like domain, Transmembrane; Glycoprotein; Signal.
KM GLYCOPROTEIN; Signal.
FT SIGNAL 1 83
FT CHAIN 84 1408 SERATE PROTEIN.
FT DOMAIN 84 1223 EXTRACELLULAR (POTENTIAL),
FT TRANSMEM 1224 1249 POTENTIAL.

```

[illegible]

```

Oy      26 CDMPCCNGCIGLPLGANGSCCECPDGFDPNCS 60
Db      803 CSDFPCRNCGICLDGG--DDDTFCBNSCHTKRCS 835

RESULT          4
NOTC_BRARE     STANDARD:    PRT:   2437 AA.
ID AC           ID AC        ID AC        ID AC
P46530;         OI-NOV-1995 (Rel. 32). Created)
DE              DE 01-NOV-1995 (Rel. 32). Last sequence update)
DR              DR 01-NOV-1995 (Rel. 32). Last annotation update)
DN              DN NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN PRECURSOR.
GC              GC NOTCH.
OS              OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC              OC Eukaryotes; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
                  Actinopterygii; Neoplatrygii; Teleostei; Euteleostei; Ostariophysi;
                  Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX              NCBI_TaxId=7955;
RN              RN [1]
RP              RP SEQUENCE FROM N.A.
RR              RR MEDLINE=94126607; PUBMED=8297791;
RX              RX Bieriemp C., Campos-Ortiga J.A.;
RA              RA "A zebrafish homologue of the Drosophila neurogenic gene Notch and
RL              RL its pattern of transcription during early embryogenesis.";
RM              RM Mech. Dev. 43:87-100(1993).
RT              RT -1 FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING
                      ENBRIO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE
                      NEURAL PLATE, NOTOCHORD AND BRAIN VESICLES.
CC              CC -1 DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PERGASTRULATION
                      STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED.
CC              CC ACCUMULATING PREDOMINANTLY IN THE PRECHORAL MESODERM AND
                      NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE
                      ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE
                      AND DIFFERENTIATING MESEDERM. ALSO PRESENT IN THE DEVELOPING
                      BRAIN AND HEAD REGIONS.
CC              CC -1 SIMILARITY: HIGH WITH OTHER NOTCH-TYPE PROTEINS.
CC              CC -1 SIMILARITY: CONTAINS 36 EGFLIKE DOMAINS.
CC              CC -1 SIMILARITY: CONTAINS 6 ANK REPEATS.
-----
CC              CC This SWISS-PROT entry is copyright. It is produced through a collaboration
                      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
                      CC the European Bioinformatics Institute. There are no restrictions on its use by
                      CC use by non-profit institutions as long as its content is in no way
                      CC modified and this statement is not removed. Usage by, or for commercial
                      CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
                      CC or send an email to licenses@isb.sib.ch).
CC              CC -----
Cc            Cc EMBL X69088: CA448831.1 ;
Cc            Cc HSSP: PD0740: IIXA.
Cc            Cc ZFIN: ZDB-GENE-990415-113; notch.
Cc            Cc InterPro: IPRO00152; -
Cc            Cc Dr InterPro: IPRO00561; -
Cc            Cc Dr InterPro: IPRO00800; -
Cc            Cc Dr InterPro: IPRO01336; -
Cc            Cc Dr InterPro: IPRO01438; -
Cc            Cc Dr InterPro: IPRO01891; -
Cc            Cc Pfam PF00080; RGRF_16.
Cc            Cc Pfam PF00023; ank_6.
Cc            Cc Pfam PF00066; notch_3.
Cc            Cc Dr PRINTS: PR00009; EGPGFG.
Cc            Cc Dr PRINTS: PR00010; EGPHLP.
Cc            Cc PROSITE: PS50089; ANK_REPEAT_4.
```



```

Db 678 DDGALNPHNGTCTDGG--VNSFTCLCPDGRFATCIS 713
RESULT 5
NTCL_HUMAN STANDARD: PRT: 2444 AA.
AC P46531:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 1 PRECURSOR (TRANSLATION-
DE ASSOCIATED NOTCH PROTEIN TAN-1) (FRAGMENT).
GN NOTCH1 OR TANI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eulheria; Primates; Carnivora; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91347367; PubMed=1831692;
RA Eilisen L.W., Bird J., West D.C., Soreng A.L., Reynolds T.C.,
RT "TAN-1, the human homolog of the Drosophila notch gene, is broken by
RT chromosomal translocations in T lymphoblastic neoplasms."
RL Cell 66:649-661(1991).
CC -1- FUNCTION: MAY BE IMPORTANT FOR NORMAL LYMPHOCYTE FUNCTION. IN
CC ALTERED FORM, MAY CONTRIBUTE TO TRANSFORMATION OR PROGRESSION
CC IN SOME T-CELL NEOPLASMS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: IN FETAL TISSUES MOST ABUNDANT IN SPLEEN,
CC BRAIN STEM AND LUNG. ALSO PRESENT IN MOST ADULT TISSUES WHERE IT
CC IS FOUND MAINLY IN LYMPHOID TISSUES.
CC -1- SIMILARITY: HIGH WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M73980; AAA60614.1; -
DR HSP: P00740; IIXA.
DR MIM: 190198; -
DR InterPro: IPR000152; -
DR InterPro: IPR000561; -
DR InterPro: IPR000800; -
DR InterPro: IPR001881; -
DR InterPro: IPR002110; -
DR Pfam: PF00008; EGF_36.
DR Pfam: PF00023; ank_6.
DR Pfam: PF00066; notch_3.
DR Prosite: PS50088; ANK REPEAT; 4.
DR Prosite: PS50297; ANK REP. REGION; 1.
DR Prosite: PS00010; ASK_HYDROXYL; 20.
DR Prosite: PS00022; EGF_1; 34.
DR Prosite: PS01186; EGF_2; 26.
DR Prosite: PS01187; EGF_CA; 18.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 18
FT CHAIN 19 >2444 POTENTIAL.
FT DOMAIN 19 1736 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 1.
FT TRANSMEM 1737 1757 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1758 >2444 POTENTIAL.
FT DOMAIN 20 58 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 59 99 EGF-LIKE 1.
FT DOMAIN 102 139 EGF-LIKE 2.
FT DOMAIN 102 139 EGF-LIKE 3.
FT DOMAIN 140 176
FT DOMAIN 178 216 EGF-LIKE 4.
FT DOMAIN 218 255 EGF-LIKE 5.
FT DOMAIN 257 293 EGF-LIKE 6.
FT DOMAIN 295 333 EGF-LIKE 7.
FT DOMAIN 335 371 EGF-LIKE 8.
FT DOMAIN 372 410 EGF-LIKE 9.
FT DOMAIN 412 450 EGF-LIKE 10.
FT DOMAIN 452 488 EGF-LIKE 11.
FT DOMAIN 490 526 EGF-LIKE 12.
FT DOMAIN 528 564 EGF-LIKE 13.
FT DOMAIN 566 601 EGF-LIKE 14.
FT DOMAIN 603 639 EGF-LIKE 15.
FT DOMAIN 641 676 EGF-LIKE 16.
FT DOMAIN 678 714 EGF-LIKE 17.
FT DOMAIN 716 751 EGF-LIKE 18.
FT DOMAIN 753 789 EGF-LIKE 19.
FT DOMAIN 791 827 EGF-LIKE 20.
FT DOMAIN 829 868 EGF-LIKE 21.
FT DOMAIN 870 906 EGF-LIKE 22.
FT DOMAIN 908 944 EGF-LIKE 23.
FT DOMAIN 946 982 EGF-LIKE 24.
FT DOMAIN 984 1020 EGF-LIKE 25.
FT DOMAIN 1022 1058 EGF-LIKE 26.
FT DOMAIN 1060 1096 EGF-LIKE 27.
FT DOMAIN 1098 1144 EGF-LIKE 28.
FT DOMAIN 1146 1182 EGF-LIKE 29.
FT DOMAIN 1184 1220 EGF-LIKE 30.
FT DOMAIN 1222 1266 EGF-LIKE 31.
FT DOMAIN 1268 1306 EGF-LIKE 32.
FT DOMAIN 1308 1347 EGF-LIKE 33.
FT DOMAIN 1349 1385 EGF-LIKE 34.
FT DOMAIN 1427 1461 EGF-LIKE 35.
FT DOMAIN 1446 1481 EGF-LIKE 36.
FT DOMAIN 1482 1523 LIN/NOTCH 1.
FT DOMAIN 1524 1563 LIN/NOTCH 2.
FT DOMAIN 1563 1597 LIN/NOTCH 3.
FT DOMAIN 1597 1631 ANK 1.
FT DOMAIN 1631 1665 ANK 2.
FT DOMAIN 1665 1699 ANK 3.
FT DOMAIN 1699 1732 ANK 4.
FT DOMAIN 1732 1766 ANK 5.
FT DOMAIN 1766 1800 POLY-ARG.
FT DOMAIN 1800 1834 POLY-VAL.
FT DOMAIN 1834 1868 POLY-ARG.
FT DOMAIN 1868 1902 POLY-PRO.
FT DOMAIN 1902 1936 POLY-ALA.
FT DOMAIN 1936 1970 POLY-GLU.
FT DOMAIN 1970 2004 POLY-GLY.
FT DOMAIN 2004 2038 POLY-GLN.
FT DOMAIN 2038 2072 POLY-PRO.
FT DOMAIN 2072 2106 BY SIMILARITY.
FT DOMAIN 2106 2140 BY SIMILARITY.
FT DOMAIN 2140 2174 BY SIMILARITY.
FT DOMAIN 2174 2208 BY SIMILARITY.
FT DOMAIN 2208 2242 BY SIMILARITY.
FT DOMAIN 2242 2276 BY SIMILARITY.
FT DOMAIN 2276 2310 BY SIMILARITY.
FT DOMAIN 2310 2344 BY SIMILARITY.
FT DOMAIN 2344 2378 BY SIMILARITY.
FT DOMAIN 2378 2412 BY SIMILARITY.
FT DOMAIN 2412 2446 BY SIMILARITY.
FT DOMAIN 2446 2480 BY SIMILARITY.
FT DOMAIN 2480 2514 BY SIMILARITY.
FT DOMAIN 2514 2548 BY SIMILARITY.
FT DOMAIN 2548 2582 BY SIMILARITY.
FT DOMAIN 2582 2616 BY SIMILARITY.
FT DOMAIN 2616 2650 BY SIMILARITY.
FT DOMAIN 2650 2684 BY SIMILARITY.
FT DOMAIN 2684 2718 BY SIMILARITY.
FT DOMAIN 2718 2752 BY SIMILARITY.
FT DOMAIN 2752 2786 BY SIMILARITY.
FT DOMAIN 2786 2820 BY SIMILARITY.
FT DOMAIN 2820 2854 BY SIMILARITY.
FT DOMAIN 2854 2888 BY SIMILARITY.
FT DOMAIN 2888 2922 BY SIMILARITY.
FT DOMAIN 2922 2956 BY SIMILARITY.
FT DOMAIN 2956 2990 BY SIMILARITY.
FT DOMAIN 2990 3024 BY SIMILARITY.
FT DOMAIN 3024 3058 BY SIMILARITY.
FT DOMAIN 3058 3092 BY SIMILARITY.
FT DOMAIN 3092 3126 BY SIMILARITY.
FT DOMAIN 3126 3160 BY SIMILARITY.
FT DOMAIN 3160 3194 BY SIMILARITY.
FT DOMAIN 3194 3228 BY SIMILARITY.
FT DOMAIN 3228 3262 BY SIMILARITY.
FT DOMAIN 3262 3296 BY SIMILARITY.
FT DOMAIN 3296 3330 BY SIMILARITY.

```

FT	DISULFID	339	350	BY SIMILARITY.	AC	001705;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
----	----------	-----	-----	----------------	----	---------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

RC	TISSUE=Gastrula;
CC	MEDLINE=93273088; PubMed=850658;
RA	Bisgrove B.W., Raff R.A.;
RT	"The Sogef III gene encodes a member of the fibropellins: EGF repeat-
RT	containing proteins that form the apical lamina of the sea urchin
RT	embryo.";
RL	Dev Biol. 157:526-538(1993).
CC	-1- FUNCTION: FORM THE APICAL LAMINA, A COMPONENT OF THE EXTRACELLULAR
CC	MATRIX.
CC	-1- SUPERCELLULAR LOCATION: EXTRACELLULAR.
CC	-1- DEVELOPMENTAL STAGE: LOW LEVELS IN UNFERTILIZED EGGS AND DURING
CC	EARLY CLEAVAGE, THEN RAPIDLY INCREASES IN ABUNDANCE BETWEEN LATE
CC	MORULA AND MESENCHYME BLASTULA STAGES TO MAXIMAL LEVELS MAINTAINED
CC	THROUGH SUBSEQUENT STAGES.
CC	-1- MISCELLANEOUS: EXPRESSED BOTH MATERNAALLY AND ZYGOTICALLY.
CC	-1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC	-1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
CC	-1- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR
CC	TO AVIDIN/SREPTAVIDIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation. It is
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@sib-sib.ch).
CC	-----
DR	EMBL; L07045; AAA30045.1; -
DR	HSSP; P00740; IIXA.
DR	InterPro; IPRO00088; -
DR	InterPro; IPRO00152; -
DR	InterPro; IPRO00561; -
DR	InterPro; IPRO00859; -
DR	InterPro; IPRO01438; -
DR	InterPro; IPRO01881; -
DR	pfam; PF01382; Avidin; 1.
DR	pfam; PF00431; CUB; 1.
DR	pfam; PF00008; EGF; 8.
DR	PRINTS; PR00010; EGFBLD.
DR	PROSITE; PS00010; ASX_HYDROXYL; 8.
DR	PROSITE; PS00022; EGF_1; 8.
DR	PROSITE; PS00577; AVIDIN; 1.
DR	PROSITE; PS01180; CUB; 1.
DR	PROSITE; PS01186; EGF_2; 7.
DR	PROSITE; PS01187; EGF_CA; 6.
KW	Biotin; EGF-like domain; Repeat; Signal; Glycoprotein;
KW	Calcium-binding.
FT	SIGNAL 1 17
FT	POTENTIAL.
FT	CHAIN 18 570
FT	FIBROPELLIN C.
FT	DOMAIN 62 175
FT	CUB.
FT	EGF-LIKE 2, CALCIIUM-BINDING (POTENTIAL).
FT	DOMAIN 214 250
FT	EGF-LIKE 4, CALCIIUM-BINDING (POTENTIAL).
FT	DOMAIN 252 288
FT	EGF-LIKE 5, CALCIIUM-BINDING (POTENTIAL).
FT	DOMAIN 290 326
FT	EGF-LIKE 6, CALCIIUM-BINDING (POTENTIAL).
FT	DOMAIN 328 364
FT	EGF-LIKE 7,
FT	DOMAIN 366 402
FT	EGF-LIKE 8, CALCIIUM-BINDING (POTENTIAL).
FT	DOMAIN 404 440
FT	AVIDIN-LIKE.
FT	DISULFID 442 570
FT	BY SIMILARITY.
FT	DISULFID 23 34
FT	BY SIMILARITY.
FT	DISULFID 28 43
FT	BY SIMILARITY.
FT	DISULFID 45 54
FT	BY SIMILARITY.
FT	DISULFID 180 191
FT	BY SIMILARITY.
FT	DISULFID 185 200
FT	BY SIMILARITY.
FT	DISULFID 202 211
FT	BY SIMILARITY.
FT	DISULFID 218 229
FT	BY SIMILARITY.
FT	DISULFID 223 238
FT	BY SIMILARITY.
FT	DISULFID 240 249
FT	BY SIMILARITY.
FT	DISULFID 256 267
FT	BY SIMILARITY.
FT	DISULFID 261 276
FT	BY SIMILARITY.
FT	DISULFID 278 287
FT	BY SIMILARITY.
FT	DISULFID 294 305
FT	BY SIMILARITY.

```

FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 316 325 BY SIMILARITY.
FT DISULFID 332 343 BY SIMILARITY.
FT DISULFID 337 352 BY SIMILARITY.
FT DISULFID 354 363 BY SIMILARITY.
FT DISULFID 370 381 BY SIMILARITY.
FT DISULFID 375 390 BY SIMILARITY.
FT DISULFID 392 401 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 428 BY SIMILARITY.
FT DISULFID 430 439 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 570 AA: 61116 MW: 8655E3E1C05E6EE CRC64;

Query Match 31.4%; Score 113.5; DB 1; Length 570;
Best Local Similarity 42.4%; Pred. No. 4.9e-05;
Matches 25; Conservative 9; Mismatches 20; Indels 5; Gaps 3;

QY 1 MKRSVAWLVGLGVPGFGKGDICDPNCGNGICLPGLAVGSPSCCPGFTDPMC 59
DB 1 MKVSLAVLL--LSIVAATYGGGE--CGSNPCBNSSVCRDG--EGTITCEQMGKIDGQNC 54

RESULT 9
ID DLK_MOUSE STANDARD; PRT; 385 AA.
AC Q09163; 007645; 062208;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DELTA-LIKE PROTEIN PRECURSOR (DLK) (PREDAPPOCYTE FACTOR 1) (PREF-1)
DE (ADIPOCYTE DIFFERENTIATION INHIBITOR PROTEIN) [CONTAINS: FETAL ANTIGEN
DE 1 (FAL)].
GN DLK1 OR DLK OR PREF1 OR SCP-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS; TISSUE=Fibroblast;
RX MEDLINE=93179372; PubMed=8095043;
RA Laborda J., Sausville E.A., Hoffman T., Notario V.;
RT "dlk, a putative mammalian homeotic gene differentially expressed in
RT small cell lung carcinoma and neuroendocrine tumor cell line.";
RL J. Biol. Chem. 268:3817-3820(1993).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=93272313; PubMed=8500166;
RA Smas C.M., Sul H.S.;
RT "Pref-1, a protein containing EGF-like repeats, inhibits adipocyte
RT differentiation.";
RL Cell 73:725-734(1993).
RN 13
RP SEQUENCE FROM N.A., AND VARIANT LYS-347 DEL.
RC TISSUE=Adrenal gland, and Placenta;
RX MEDLINE=95226449; PubMed=7711066;
RA Lee Y.L., Helman L., Hoffman T., Laborda J.;
RT "dlk, pg2 and Pref-1 mRNAs encode similar proteins belonging to the
RT EGF-like superfamily. Identification of polymorphic variants of this
RT RNA.";
RL Biochim. Biophys. Acta 1261:223-232(1995).
RN 14
RP SEQUENCE FROM N.A.
RA Maruyama K., Nishijima S., Kuromitsu S., Ichikawa A., Masuda E.,
RA Yakemoto T., Kodama H., Kawashima H.;
RN Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE OF 1-8 FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=94325292; PubMed=7519443;

RA Smas C.M., Green D., Sul H.S.;
RT "Structural characterization and alternate splicing of the gene
RT encoding the preadipocyte EGF-like protein pref-1.";
RL Biochemistry 33:9257-9265(1994).
RN 16
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=97234561; PubMed=9116998;
RA Krogh T.N., Bachmann E., Teisner B., Skjodt K., Hoegrup P.;
RT "Glycosylation analysis and protein structure determination of murine
RT fetal antigen 1 (mFAL) -- the circulating gene product of the
RT delta-like protein (dlk), preadipocyte factor 1 (Pref-1) and
RT stromal-cell-derived protein 1 (SCP-1) cDNAs.";
RL Eur. J. Biochem. 244:334-342(1997).
CC -1- FUNCTION: MAY HAVE A ROLE IN NEUROENDOCRINE DIFFERENTIATION.
CC -1- INHIBITS ADIPOCYTE DIFFERENTIATION.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: A (SHOWN HERE), B, C,
CC C2, D AND D2: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 212171; CAA78162.1; -
DR EMBL: U15980; AAB60495.1; -
DR EMBL: L12721; AAA37175.1; -
DR EMBL: S71340; -; NOT ANNOTATED_CDS.
DR EMBL: D16847; BAA04121.1; -
DR HSRP: P00740; IIXA
DR Glycosultemp; Q09163; -
DR MGD; MG194900; DLK1.
DR Interpro; IPR000152; -
DR Interpro; IPR000561; -
DR Pfam; PF00008; EGF; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 6.
KW Antigen; Transmembrane; EGF-like domain; Repeat; Glycoprotein; Signal;
KW Alternative splicing.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 385 DELTA-LIKE PROTEIN.
FT CHAIN 24 305 FETAL ANTIGEN 1.
FT DOMAIN 24 305 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 306 329 POTENTIAL.
FT DOMAIN 330 385 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 353 385 EGF-LIKE 1.
FT DOMAIN 353 385 EGF-LIKE 2.
FT DOMAIN 88 125 EGF-LIKE 3.
FT DOMAIN 127 168 EGF-LIKE 4.
FT DOMAIN 172 208 EGF-LIKE 5.
FT DOMAIN 210 247 EGF-LIKE 6.
FT DISULFID 26 37 BY SIMILARITY.
FT DISULFID 30 43 BY SIMILARITY.
FT DISULFID 45 54 BY SIMILARITY.
FT DISULFID 57 68 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 76 85 BY SIMILARITY.
FT DISULFID 92 103 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 156 BY SIMILARITY.
FT DISULFID 158 167 BY SIMILARITY.
FT DISULFID 176 187 BY SIMILARITY.
FT DISULFID 181 196 BY SIMILARITY.
FT DISULFID 198 207 BY SIMILARITY.

```

```

FT  DISULFID  214  225  BY SIMILARITY.
FT  DISULFID  219  235  BY SIMILARITY.
FT  DISULFID  237  246  BY SIMILARITY.
FT  CARBOHYD  94  94  O-LINKED.
FT  CARBOHYD  100  100  N-LINKED (GLCNAC. . .)
FT  CARBOHYD  165  165  N-LINKED (GLCNAC. . .)
FT  CARBOHYD  174  174  N-LINKED (GLCNAC. . .)
FT  CARBOHYD  216  216  O-LINKED.
FT  CARBOHYD  224  224  O-LINKED.
FT  CARBOHYD  258  258  O-LINKED.
FT  CARBOHYD  267  267  O-LINKED (PARTIAL).
FT  CARBOHYD  271  271  O-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD  295  295  MISSING (IN ISOFORM D).
FT  VARSPLC  211  303  MISSING (IN ISOFORM D2).
FT  VARSPLC  231  281  MISSING (IN ISOFORM B).
FT  VARSPLC  231  303  MISSING (IN ISOFORM C).
FT  VARSPLC  231  305  MISSING (IN ISOFORM C2).
FT  VARIANT  347  347  MISSING.
FT  CONFLICT  250  250  R -> P (IN REF. 2).
FT  CONFLICT  320  385  VIGTVAIVFINKCTEWNSNLRNHPFRKKKNILOYNGGEE
                          LAVNIIPEKIDMTFNKNGDEPT -> CMAPWPSSTST
                          AKRGPTCATTTTCARRRTSCCSTTAARSRKSISSPRRLT
                          (IN REF. 2).
                          TF -> WL (IN REF. 4).
FT  CONFLICT  344  345  E79864FEA5AF4FPI CRC64;
SO  SEQUENCE  385 AA; 41320 MM; E79864FEA5AF4FPI CRC64;

Query Match      31.2%; Score 113; DB 1; Length 385;
Best Local Similarity 52.6%; Pred. No. 4e-05;
Matches 20; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

Oy  26 CDPNPCENGICLPGIAGVSFSCPCPDGFTDPCSSVY 63
Db  176 CTPNCPENGVCVD--IGGDFRCRCRAGFYDTCSPV 211

RESULT 10
DL1L_HUMAN  STANDARD; PRT; 723 AA.
ID  DL1L_HUMAN
AC  000548;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  DELTA-LIKE PROTEIN 1 PRECURSOR (DROSOPHILA DELTA HOMOLOG 1) (DELTA1)
DE  (H-DELTA-1).
GN  DL1L
OS  Homo sapiens (Human).
OC  Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99180765; PubMed=10079256;
RA  Gray G.E., Mann R.S., Mitsiadis E., Carcangiu M.L.,
RA  Banks A., Lehman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
PT  "Human ligands of the Notch receptor.";
RL  Am. J. Pathol. 154:785-794(1999).
CC  -1- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
CC  MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
CC  UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY
CC  SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC  -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC  -1- SIMILARITY: BELONGS TO THE DELTA/SERRAT/JAGGED FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).

```

```

CC  EMBL: AF003522; AB61286.1; -.
DR  HSSP: P00740; IIXA.
DR  InterPro: IPR000152; -.
DR  InterPro: IPR000561; -.
DR  InterPro: IPR001438; -.
DR  InterPro: IPR001774; -.
DR  InterPro: IPR001881; -.
DR  Pfam: PF01414; DSL; 1.
DR  Pfam: PF00008; EGF; 8.
DR  PRINTS: PR00010; EGFBL00D.
DR  PROSITE: PS00010; ASX_HYDROXYL, 3.
DR  PROSITE: PS00022; EGF_1; 8.
DR  PROSITE: PS01186; EGF_2; 8.
DR  PROSITE: PS01187; EGF_CA_1.
KW  Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
KW  Differentiation; Glycoprotein.
FT  SIGNAL  1  17  DELTA-LIKE PROTEIN 1.
FT  CHAIN  18  723  EXTRACELLULAR (POTENTIAL).
FT  DOMAIN  18  545  POTENTIAL.
FT  TRAMEM  546  568  CYTOPLASMIC (POTENTIAL).
FT  DOMAIN  569  723  EGF-LIKE 1.
FT  DOMAIN  257  285  EGF-LIKE 2.
FT  DOMAIN  325  325  EGF-LIKE 3.
FT  DOMAIN  332  363  EGF-LIKE 4, CALCIUM BINDING (POTENTIAL).
FT  DOMAIN  370  402  EGF-LIKE 5.
FT  DOMAIN  409  440  EGF-LIKE 6.
FT  DOMAIN  447  478  EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT  DOMAIN  485  516  EGF-LIKE 8.
FT  DISULFID  226  237  BY SIMILARITY.
FT  DISULFID  230  243  BY SIMILARITY.
FT  DISULFID  245  254  BY SIMILARITY.
FT  DISULFID  257  268  BY SIMILARITY.
FT  DISULFID  263  274  BY SIMILARITY.
FT  DISULFID  276  285  BY SIMILARITY.
FT  DISULFID  292  304  BY SIMILARITY.
FT  DISULFID  298  314  BY SIMILARITY.
FT  DISULFID  316  325  BY SIMILARITY.
FT  DISULFID  332  343  BY SIMILARITY.
FT  DISULFID  337  352  BY SIMILARITY.
FT  DISULFID  354  363  BY SIMILARITY.
FT  DISULFID  370  381  BY SIMILARITY.
FT  DISULFID  375  391  BY SIMILARITY.
FT  DISULFID  393  402  BY SIMILARITY.
FT  DISULFID  409  420  BY SIMILARITY.
FT  DISULFID  414  429  BY SIMILARITY.
FT  DISULFID  431  440  BY SIMILARITY.
FT  DISULFID  447  467  BY SIMILARITY.
FT  DISULFID  469  478  BY SIMILARITY.
FT  DISULFID  485  496  BY SIMILARITY.
FT  DISULFID  490  505  BY SIMILARITY.
FT  DISULFID  507  516  BY SIMILARITY.
FT  CARBOHYD  477  477  N-LINKED (GLCNAC. . .) (POTENTIAL).
SO  SEQUENCE  723 AA; 77956 MM; B4EC455F3A32A12B CRC64;

Query Match      31.2%; Score 113; DB 1; Length 723;
Best Local Similarity 52.5%; Pred. No. 6.7e-05;
Matches 21; Conservative 3; Mismatches 14; Indels 2; Gaps 1;

Oy  24 DICDPNCPENGICLPGIAGVSFSCPCPDGFTDPCSSVY 63
Db  445 DDCASPCANGTCRCD--VNDSCICPCPGYTGNCSAPV 482

RESULT 11
NOTC_XENLA  STANDARD; PRT; 2524 AA.
ID  NOTC_XENLA
AC  F21783;
DT  01-MAY-1991 (Rel. 18, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)

```



```

FT DISULFID 574 588 BY SIMILARITY.
FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 617 BY SIMILARITY.
FT DISULFID 611 626 BY SIMILARITY.
FT DISULFID 628 637 BY SIMILARITY.
FT DISULFID 644 654 BY SIMILARITY.
FT DISULFID 649 663 BY SIMILARITY.
FT DISULFID 665 674 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
FT DISULFID 686 701 BY SIMILARITY.
FT DISULFID 703 712 BY SIMILARITY.
FT DISULFID 719 729 BY SIMILARITY.
FT DISULFID 724 738 BY SIMILARITY.
FT DISULFID 740 749 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 761 776 BY SIMILARITY.
FT DISULFID 778 787 BY SIMILARITY.
FT DISULFID 794 805 BY SIMILARITY.
FT DISULFID 799 814 BY SIMILARITY.
FT DISULFID 816 825 BY SIMILARITY.
FT DISULFID 832 843 BY SIMILARITY.
FT DISULFID 837 854 BY SIMILARITY.
FT DISULFID 856 865 BY SIMILARITY.
FT DISULFID 872 883 BY SIMILARITY.
FT DISULFID 877 892 BY SIMILARITY.
FT DISULFID 894 903 BY SIMILARITY.
FT DISULFID 910 921 BY SIMILARITY.
FT DISULFID 915 930 BY SIMILARITY.
FT DISULFID 932 941 BY SIMILARITY.
FT DISULFID 966 997 BY SIMILARITY.
FT DISULFID 991 1006 BY SIMILARITY.
FT DISULFID 1008 1017 BY SIMILARITY.
FT DISULFID 1024 1035 BY SIMILARITY.
FT DISULFID 1029 1044 BY SIMILARITY.
FT DISULFID 1046 1055 BY SIMILARITY.
FT DISULFID 1062 1073 BY SIMILARITY.
FT DISULFID 1067 1082 BY SIMILARITY.
FT DISULFID 1084 1093 BY SIMILARITY.
FT DISULFID 1100 1121 BY SIMILARITY.
FT DISULFID 1115 1130 BY SIMILARITY.
FT DISULFID 1132 1141 BY SIMILARITY.
FT DISULFID 1148 1159 BY SIMILARITY.
FT DISULFID 1153 1168 BY SIMILARITY.
FT DISULFID 1170 1179 BY SIMILARITY.
FT DISULFID 1186 1197 BY SIMILARITY.
FT DISULFID 1191 1206 BY SIMILARITY.
FT DISULFID 1208 1217 BY SIMILARITY.
FT DISULFID 1224 1243 BY SIMILARITY.
FT DISULFID 1237 1252 BY SIMILARITY.

```

Query Match 31.2%; Score 113; DB 1; Length 2524;
 Best Local Similarity 52.8%; Pred. No. 0.00019;
 Matches 19; Conservative 1; Mismatches 14; Indels 2; Gaps 1;

Oy 24 DICDPKPCNGGICLPGIAVSFSCPCPDGRTDPC 59
 Db 908 DDCQPMCHNGSSGSDG--IMKPCNCPAGRGPKC 941

```

RESULT 12
DLK_HUMAN STANDARD; PRT; 383 AA.
ID P80370;
AC 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE DELTA-LIKE PROTEIN PRECURSOR (DLK) (CONTAINS: FETAL ANTIGEN 1 (FA1)).
GN DLK1 OR DLK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
FT [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE-Adrenal gland;
RX MEDLINE=93179372; PubMed=8095043;
RA Laborda J., Sausville E.A., Hoffman T., Notario V.;
RT "dlk", a putative mammalian homeotic gene differentially expressed in
RT small cell lung carcinoma and neuroendocrine tumor cell line";
RL J. Biol. Chem. 268:3817-3820(1993).
RN [12]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta, and Adrenal gland;
RX MEDLINE=95226449; PubMed=7711066;
RA Lee Y.L., Helman L., Hoffman T., Laborda J.;
RT "dlk pg2 and Pref-1 mRNAs encode similar proteins belonging to the
RT EGF-like superfamily. Identification of polymorphic variants of this
RT RNA.";
RL Biochim. Biophys. Acta 1261:223-232(1995).
RN [13]
RP SEQUENCE OF 24-383.
RC TISSUE-Amniotic fluid;
RX MEDLINE=95010145; PubMed=7925474;
RA Jensen C.H., Krogh T.N., Hoelrup P., Clausen P.P., Skjodt K.,
RA Larsson L.-J., Enghild J.J., Teisner B.;
RT "Protein structure of fetal antigen 1 (FA1). A novel circulating
RT human epidermal-growth-factor-like protein expressed in
RT neuroendocrine tumors and its relation to the gene products of dlk
RT and pg2.";
RL Eur. J. Biochem. 225:83-92(1994).
RN [14]
RP SEQUENCE OF 24-60.
RC TISSUE-Amniotic fluid;
RX MEDLINE=93273893; PubMed=8501199;
RA Jensen C.H., Teisner B., Hoelrup P., Rasmussen H.B., Madsen O.D.,
RA Nielsen B., Skjodt K.;
RT "Studies on the isolation, structural analysis and tissue localization
RT of fetal antigen 1 and its relation to a human adrenal-specific cDNA,
RT pg2.";
RL Hum. Reprod. 8:635-641(1993).
RN [15]
RP Hum. Reprod. 8:635-641(1993).
RC "FUNCTION: MAY HAVE A ROLE IN NEUROENDOCRINE DIFFERENTIATION.
CC -I- SUBUNIT: MONOMER.
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: A LONG FORM (SHOWN
CC HERE) AND A SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -I- TISSUE SPECIFICITY: FA1 IS FOUND WITHIN THE STROMAL CELLS IN CLOSE
CC CONTACT TO THE VASCULAR STRUCTURE OF PLACENTAL VILLI, YOLK SAC OF
CC FETAL LIVER, ADRENAL CORTEX AND PANCREAS AND IN THE BETA CELLS OF
CC THE ISLETS OF LANGERHANS IN THE ADULT PANCREAS. FOUND ALSO IN SOME
CC FORMS OF NEUROENDOCRINE LONG TUMOR TISSUE.
CC -I- PTM: N- AND O-GLYCOSYLATED.
CC -I- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z12173; CAAT8163.1; -
DR EMBL: U15979; AAA75364.1; -
DR EMBL: U15981; AAA75365.1; -
DR HSP: P00750; ITPG.
DR MIM: 176290; -
DR InterPro: IPR000152; -
DR InterPro: IPR000561; -
DR Pfam: PF00008; EGF_6.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00022; EGF_1; 5.
DR PROSITE: PS01186; EGF_2; 5.
KW Antigen; Transmembrane; EGF-like domain; Repeat; Glycoprotein; Signal;
KW Alternative splicing.
FT SIGNAL 1 23
FT CHAIN 24 383 DELTA-LIKE PROTEIN.

```

```

FT CHAIN 24 303 FETAL ANTIGEN 1
FT DOMAIN 24 303 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 304 327 POTENTIAL.
FT DOMAIN 328 353 CYTOSOLASMIC (POTENTIAL).
FT DOMAIN 24 55 EGF-LIKE 1.
FT DOMAIN 53 86 EGF-LIKE 2.
FT DOMAIN 88 125 EGF-LIKE 3.
FT DOMAIN 127 168 EGF-LIKE 4.
FT DOMAIN 170 206 EGF-LIKE 5.
FT DOMAIN 208 245 EGF-LIKE 6.
FT DISULFID 26 37 BY SIMILARITY.
FT DISULFID 30 43 BY SIMILARITY.
FT DISULFID 45 54 BY SIMILARITY.
FT DISULFID 57 68 BY SIMILARITY.
FT DISULFID 74 85 BY SIMILARITY.
FT DISULFID 92 103 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 156 BY SIMILARITY.
FT DISULFID 158 167 BY SIMILARITY.
FT DISULFID 174 185 BY SIMILARITY.
FT DISULFID 179 194 BY SIMILARITY.
FT DISULFID 196 205 BY SIMILARITY.
FT DISULFID 212 223 BY SIMILARITY.
FT DISULFID 217 233 BY SIMILARITY.
FT DISULFID 235 244 BY SIMILARITY.
FT CARBOHYD 94 94 O-LINKED.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .).
FT CARBOHYD 143 143 O-LINKED.
FT CARBOHYD 163 163 O-LINKED (PARTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (PARTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (PARTIAL).
FT CARBOHYD 214 214 O-LINKED.
FT CARBOHYD 222 222 O-LINKED (PARTIAL).
FT CARBOHYD 251 251 O-LINKED (PARTIAL).
FT CARBOHYD 260 260 O-LINKED (PARTIAL).
FT CARBOHYD 260 260 MISSING (IN SHORT ISOROM).
FT VARSPLIC 229 301 MISSING (IN CLONE HDLRAAG).
FT VARIANT 347 347 /FTID-VAR_002274.
FT CONFLICT 46 47 QP -> HV (IN REF. 1).
FT CONFLICT 108 108 G -> D (IN REF. 3).
SQ SEQUENCE 383 AA; 41143 MW; 205506E355B4415B CRC64;

Query Match 30.9% Score 112; DB 1; Length 383;
Best Local Similarity 52.6%; Pred. NO. 5.1e-05;
Matches 20; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

QY 26 CDPNCEGICLPLGLAVSFCSCPCDGFPTDPCSSVY 63
DB 174 CTPNCEGVCYCD-IGDFRCPCAGFTDKCSRPV 209

RESULT 13
ID FA10_CHICK STANDARD; PRT: 475 AA.
AC P25155;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE COAGULATION FACTOR X PRECURSOR (Ec 3.4.21.6) (STUART FACTOR)
DE (VIRUS ACTIVATING PROTEASE) (VAP).
GN FX.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Chorioallantoic membrane;

```



```

FT DISULFID 374 390 BY SIMILARITY.
FT DISULFID 392 401 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 428 BY SIMILARITY.
FT DISULFID 430 439 BY SIMILARITY.
FT DISULFID 446 466 BY SIMILARITY.
FT DISULFID 468 477 BY SIMILARITY.
FT DISULFID 484 495 BY SIMILARITY.
FT DISULFID 489 504 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT CARBOHYD 476 476 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 722 AA: 78448 MW: 95581B56DCE950 CRC64;

Query Match 30.1%; Score 109; DB 1; Length 722;
Best Local Similarity 50.0%; Pred. No. 0.00018;
Matches 20; Conservative 4; Mismatches 14; Indels 2; Gaps 1;

OY 24 DICDPNCEGICLPGLAVGSFSCPCDGTDPNCSV 63
DB 444 DDCASSPCANGTCRD--SVNDFSCTCPGYTGKNCAPV 481

RESULT 15
ID DILL RAT STANDARD: PRT: 714 AA.
AC P97677:
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE DELTA-LIKE PROTEIN 1 PRECURSOR (DROSOPHILA DELTA HOMOLOG 1) (DELTA1).
GN DILL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Disiblo G., Hebshi L., Boulter J., Weinmaster G.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
CC MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
CC UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY
CC SIMILARITY).
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC - SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC - SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U78889; AAB37343.1; -
DR HSSP; P00740; 11XA.
DR InterPro; IPR000152; -
DR InterPro; IPR000561; -
DR InterPro; IPR001438; -
DR InterPro; IPR001774; -
DR InterPro; IPR001881; -
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 7.
DR PRINTS; PRO0010; EGFBL00D.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF 1; 8.
DR PROSITE; PS01186; EGF 2; 8.
DR PROSITE; PS01187; EGF CA; 2.
KW Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
KW Differentiation; Glycoprotein.
FT SIGNAL 1 17

```

```

FT CHAIN 18 714 DELTA-LIKE PROTEIN 1.
FT DOMAIN 18 537 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 538 560 POTENTIAL.
FT DOMAIN 561 714 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 225 253 EGF-LIKE 1.
FT DOMAIN 256 284 EGF-LIKE 2.
FT DOMAIN 291 324 EGF-LIKE 3.
FT DOMAIN 331 362 EGF-LIKE 4.
FT DOMAIN 369 401 EGF-LIKE 5.
FT DOMAIN 408 439 EGF-LIKE 6.
FT DOMAIN 446 477 EGF-LIKE 7.
FT DOMAIN 484 515 EGF-LIKE 8.
FT DISULFID 225 236 BY SIMILARITY.
FT DISULFID 229 242 BY SIMILARITY.
FT DISULFID 244 253 BY SIMILARITY.
FT DISULFID 256 267 BY SIMILARITY.
FT DISULFID 262 273 BY SIMILARITY.
FT DISULFID 275 284 BY SIMILARITY.
FT DISULFID 291 303 BY SIMILARITY.
FT DISULFID 297 313 BY SIMILARITY.
FT DISULFID 315 324 BY SIMILARITY.
FT DISULFID 331 342 BY SIMILARITY.
FT DISULFID 336 351 BY SIMILARITY.
FT DISULFID 353 362 BY SIMILARITY.
FT DISULFID 369 380 BY SIMILARITY.
FT DISULFID 374 390 BY SIMILARITY.
FT DISULFID 392 401 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 428 BY SIMILARITY.
FT DISULFID 430 439 BY SIMILARITY.
FT DISULFID 446 466 BY SIMILARITY.
FT DISULFID 468 477 BY SIMILARITY.
FT DISULFID 484 495 BY SIMILARITY.
FT DISULFID 489 504 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT CARBOHYD 476 476 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 714 AA: 77378 MW: 488E227BBA27E CRC64;

Query Match 29.8%; Score 108; DB 1; Length 714;
Best Local Similarity 50.0%; Pred. No. 0.00023;
Matches 20; Conservative 4; Mismatches 14; Indels 2; Gaps 1;

OY 24 DICDPNCEGICLPGLAVGSFSCPCDGTDPNCSV 63
DB 444 DDCASSPCANGTCRD--SVNDFSCTCPGYTGKNCAPV 481

```

Search completed: May 23, 2001, 06:18:57
Job time: 361 sec

Result No.	Score	Query Match	Length	DB	ID	Description
1	355	98.1	480	1	043854	043854 homo sapien
2	304	84.0	480	11	0354754	0354754 mus musculus
3	130	35.9	2319	11	09R172	09R172 rattus norv
4	126	34.8	2281	4	09PUL3	09PUL3 homo sapien
5	126	34.8	2321	4	09V6L8	09V6L8 homo sapien
6	126	34.8	2321	4	09UM47	09UM47 homo sapien
7	124	34.3	406	5	025059	025059 heliocardia
8	121	33.4	1404	5	09VB65	09VB65 drosophila
9	117	32.3	585	5	09OUE2	09OUE2 tribolium c
10	114	31.5	529	5	025058	025058 heliocardia
11	114	31.5	2447	13	013149	013149 figu rubrifi
12	113	31.2	383	11	070534	070534 rattus norv
13	113	31.2	383	11	062779	062779 rattus norv
14	113	31.2	723	4	09OUV2	09OUV2 homo sapien
15	113	31.2	723	4	09NU41	09NU41 homo sapien
16	113	31.2	1212	13	042347	042347 gallus gal
17	113	31.2	1254	13	09YHU2	09YHU2 brachydanila
18	113	31.2	3666	5	09V6N8	09V6N8 drosophila
19	112	30.9	2653	5	025253	025253 lucilia cup

20	111.5	30.8	1075	5	08NC90	Q9nc90 strongioce
21	111	30.7	717	13	P87357	P87357 brachydanto
22	111	30.7	802	13	057462	057462 brachydanto
23	110	30.4	308	6	046370	046370 bos taurus
24	110	30.4	3671	5	020911	020911 caenorhabdi
25	109	30.1	138	6	028994	028994 sus scrofa
26	109	30.1	1200	4	Q9Y6P9	Q9Y6P9 homo sapien
27	109	30.1	1202	11	P97607	P97607 rattus norvi
28	109	30.1	1223	4	Q9UE17	Q9UE17 homo sapien
29	109	30.1	1238	4	Q9Y6D0	Q9Y6D0 homo sapien
30	109	30.1	1238	4	Q9Y219	Q9Y219 homo sapien
31	109	30.1	1238	4	Q9UNK8	Q9UNK8 homo sapien
32	109	30.1	1238	4	Q9UE99	Q9UE99 homo sapien
33	109	30.1	1247	11	Q9OYE5	Q9OYE5 mus muscullu
34	109	30.1	2471	11	Q9OM30	Q9OM30 rattus sp.
35	108	29.8	482	11	063207	063207 rattus norv
36	108	29.8	752	13	042374	042374 brachydanto
37	108	29.8	2352	5	061240	061240 halocynthia
38	108	29.8	2331	5	016004	016004 lytechinus
39	107	29.6	861	11	Q9QW58	Q9QW58 mus sp. mot
40	107	29.6	1203	11	Q06008	Q06008 mus muscullu
41	107	29.6	2470	11	Q35516	Q35516 mus muscullu
42	107	29.6	2634	5	Q9M4T8	Q9M4T8 drosophila
43	107	29.6	2704	5	097458	097458 drosophila
44	106	29.3	589	11	088671	088671 rattus norv
45	105.5	29.1	456	4	Q14316	Q14316 homo sapien

ALIGNMENTS

043854	PRELIMINARY;	PRT;	480 AA.
AC	043854; 043855;		
DT	01-JUN-1998 (TREMBLrel, 06, Created)		
DT	01-JUN-1998 (TREMBLrel, 06, Last sequence update)		
DT	01-OCT-2000 (TREMBLrel, 15, Last annotation update)		
DE	INTEGRIN-BINDING PROTEIN DELT PRECURSOR.		
DE	DELT.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.		
RC	TISSUE=EMBRYONIC LUNG;		
RX	Medline=98083109; Pubmed=9420328;		
RA	Hidal C., Zupancic T.J., Penta K., Mikhail A., Kawana M.,		
RA	Quertemous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,		
RA	Auerbach R., Hogan B.L.M., Snodgrass R., Quertemous T.;		
RT	"Cloning and characterization of developmental endothelial locus-1: an		
RT	embryonic endothelial cell protein that binds the alphabeta3 Integrin		
RT	receptor.";		
RL	Genes Dev. 12:21-33(1998).		
CC	-1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH		
CC	INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS		
CC	FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN		
CC	REGULATION OF VASCULAR MORPHOGENESIS OF REMODELING IN EMBRYONIC		
CC	DEVELOPMENT.		
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.		
CC	-1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND		
CC	SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.		
CC	-1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.		
CC	-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.		
DR	EMBL; U70312; AAC02646.1; -;		
DR	EMBL; U70313; AAC02649.1; -;		
DR	HSSP; P00740; 11XA.		
DR	INTERPRO: IPR000152; -;		
DR	INTERPRO: IPR000421; -;		
DR	INTERPRO: IPR000561; -;		
DR	INTERPRO: IPR000742; -;		
DR	INTERPRO: IPR001881; -;		

DR PFAM; PF00008; EGF_3.
 DR PFAM; PF00754; F5_F8_type_C; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01285; FAS8C_2; 2.
 DR EGF-like domain; Alternative splicing; Signal; Developmental protein;
 KW EGF-like domain; Alternative splicing; Signal; Developmental protein;
 RE Cell adhesion; Repeat; Vascularization.
 FT SIGNAL 1 16
 FT CHAIN 17 480 INTEGRIN-BINDING PROTEIN DEL1.
 FT DOMAIN 26 59 EGF-LIKE 1.
 FT DOMAIN 78 116 EGF-LIKE 2.
 FT DOMAIN 123 154 EGF-LIKE 3.
 FT DOMAIN 161 311 F5/8 TYPE C 1.
 FT DOMAIN 322 473 F5/8 TYPE C 2.
 FT SITE 96 98 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 26 37 BY SIMILARITY.
 FT DISULFID 31 48 BY SIMILARITY.
 FT DISULFID 50 59 BY SIMILARITY.
 FT DISULFID 78 89 BY SIMILARITY.
 FT DISULFID 83 105 BY SIMILARITY.
 FT DISULFID 107 116 BY SIMILARITY.
 FT DISULFID 158 314 BY SIMILARITY.
 FT DISULFID 301 305 BY SIMILARITY.
 FT DISULFID 319 476 BY SIMILARITY.
 FT VARSPIC 66 66 A -> G (IN SHORT ISOFORM).
 FT VARSPIC 67 76 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 480 AA; 53765 MW; F7171E23A309FD8 CRC64;

Query Match 98.1%; Score 355; DB 4; Length 480;
 Best Local Similarity 98.5%; Pred. No. 6e-35;
 Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKRSVAWLVGLSLGVPQFGKIDICDPNCPENGICLPGGLAVGSFSCPCPGFTDPNCS 60
 DB 1 MKRSVAWLVGLSLGVPQFGKIDICDPNCPENGICLPGGLAVGSFSCPCPGFTDPNCS 60
 QY 61 SVEEV 65
 DB 61 SVEEV 65

RESULT 2
 ID 035474 PRELIMINARY; PRT; 480 AA.
 AC 035474: 035475;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.
 GN EDL3 OR DEL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
 RC TISSUE-EMBRO;
 RX MEDLINE=98083109; PubMed=9420328;
 RA Hidaï C., Zupancic T., Penta K., Mikhail A., Kawana M.,
 RA Quertemous E.E., Aoka Y., Fukagawa M., Matsui Y., Platka D.,
 RA Auerbach R., Hogan B.L.M., Snodgrass R., Quertemous T.;
 RT "Cloning and characterization of developmental endothelial locus-1: an
 RT embryonic endothelial cell protein that binds the alphavbeta3 integrin
 RT receptor";
 RL Genes Dev. 12:21-33(1998).
 CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
 CC INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
 CC FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
 CC REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC

CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND
 CC SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
 CC CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP
 CC OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
 CC NEURONS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7, AFTER DAY
 CC 15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME
 CC OF BIRTH.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC EMBL; AF031524; AAB86585.1; -;
 CC EMBL; AF031525; AAB86586.1; -;
 CC HSSP; P00740; 1EDM.
 CC MGD; MGI:1329025; Edl13.
 CC INTERPRO: IPR000152; -;
 CC INTERPRO: IPR000421; -;
 CC INTERPRO: IPR000561; -;
 CC INTERPRO: IPR000742; -;
 CC INTERPRO: IPR01881; -;
 CC PFAM; PF00008; EGF_3.
 CC PFAM; PF00754; F5_F8_type_C; 2.
 CC PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 CC PROSITE; PS00022; EGF_1; UNKNOWN_3.
 CC PROSITE; PS01187; EGF_CA; 1.
 CC PROSITE; PS01285; FAS8C_1; 2.
 CC PROSITE; PS01285; FAS8C_2; 2.
 CC EGF-like domain; Alternative splicing; Signal; Developmental protein;
 KW EGF-like domain; Alternative splicing; Signal; Developmental protein;
 RE Cell adhesion; Repeat; Vascularization.
 FT SIGNAL 1 16
 FT CHAIN 17 480 INTEGRIN-BINDING PROTEIN DEL1.
 FT DOMAIN 26 59 EGF-LIKE 1.
 FT DOMAIN 78 116 EGF-LIKE 2.
 FT DOMAIN 123 154 EGF-LIKE 3.
 FT DOMAIN 161 311 F5/8 TYPE C 1.
 FT DOMAIN 322 473 F5/8 TYPE C 2.
 FT SITE 96 98 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 26 37 BY SIMILARITY.
 FT DISULFID 31 48 BY SIMILARITY.
 FT DISULFID 50 59 BY SIMILARITY.
 FT DISULFID 78 89 BY SIMILARITY.
 FT DISULFID 83 105 BY SIMILARITY.
 FT DISULFID 107 116 BY SIMILARITY.
 FT DISULFID 158 314 BY SIMILARITY.
 FT DISULFID 301 305 BY SIMILARITY.
 FT DISULFID 319 476 BY SIMILARITY.
 FT VARSPIC 218 221 INFO -> VTG (IN SHORT ISOFORM).
 FT VARSPIC 222 480 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 480 AA; 53740 MW; 4CD91EF9261714D CRC64;

Query Match 84.0%; Score 304; DB 11; Length 480;
 Best Local Similarity 84.6%; Pred. No. 8.5e-29;
 Matches 55; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKRSVAWLVGLSLGVPQFGKIDICDPNCPENGICLPGGLAVGSFSCPCPGFTDPNCS 60
 DB 1 MKRSVAWLVGLSLGVPQFGKIDICDPNCPENGICLPGGLAVGSFSCPCPGFTDPNCS 60
 QY 61 SVEEV 65
 DB 61 SVEEV 65

RESULT 3
 ID 09R172 PRELIMINARY; PRT; 2319 AA.
 AC 09R172:
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

```

DE 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE NOTCH 3 PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattunians T., Boulter J., Weinmaster G., Schanen N.C.;
RT "Rattus norvegicus mRNA for Notch 3."
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF164486; AAD4653.1;
DR HSSP; P00740; 1EDM.
DR INTERPRO; IPR000152;
DR INTERPRO; IPR000561;
DR INTERPRO; IPR000742;
DR INTERPRO; IPR000800;
DR INTERPRO; IPR001438;
DR INTERPRO; IPR001881;
DR INTERPRO; IPR002049;
DR INTERPRO; IPR002110;
DR PFAM; PF00008; EGF_33.
DR PFAM; PF00023; ank; 6.
DR PRINTS; PR00066; notch; 3.
DR PRINTS; PR00010; EGFRLD.
DR PRINTS; PR00011; EGFAMININ.
DR PROSITE; PS00010; ASX_HYDROXYL; 18.
DR PROSITE; PS00022; EGF_1; UNKNOWN_32.
DR PROSITE; PS01186; EGF_2; 25.
DR PROSITE; PS01187; EGF_CA; 16.
SQ SEQUENCE 2319 AA; 244358 MW; 2FE9D14D15A32D44 CRC64;

Query Match 35.98; Score 130; DB 11; Length 2319;
Best Local Similarity 58.58; Pred. No. 4.1e-07;
Matches 24; Conservative 2; Mismatches 13; Indels 2; Gaps 1;

OY 24 DCDPNCENGICLPLGLAVGSCFCECPDGFDPNCSSVVE 64
DB 853 DCDPNCPLNGSCQDQ--VGFSFCSCLPFGAGPRCARDVD 891
ID 090PL3 PRELIMINARY; PRT; 2281 AA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Ganes J., Dangnan L., Poundstone P.,
RA Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
RA Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Krommiller B., Arellano A.,
RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S.,
RA Carrano A.V.;
RT "Sequence analysis of an ~1.5 Mb olfactory receptor (OLFR) cluster in
RT 19p13.1."
RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC004663; AAC15789.1;
DR HSSP; P00740; 1IXA.
DR INTERPRO; IPR000083;
DR INTERPRO; IPR000152;
DR INTERPRO; IPR000561;

```

```

DR INTERPRO; IPR000742;
DR INTERPRO; IPR000800;
DR INTERPRO; IPR001010;
DR INTERPRO; IPR001438;
DR INTERPRO; IPR001762;
DR INTERPRO; IPR001881;
DR INTERPRO; IPR001947;
DR INTERPRO; IPR002049;
DR INTERPRO; IPR002110;
DR PFAM; PF00008; EGF_34.
DR PFAM; PF00023; ank; 6.
DR PFAM; PF00066; notch; 3.
DR PRINTS; PR00010; EGFRLD.
DR PRINTS; PR00011; EGFAMININ.
DR PRINTS; PR00012; ENTPEI.
DR PRINTS; PR00286; CHAREDTOXIN.
DR PRINTS; PR00287; THIONIN.
DR PRINTS; PR00289; DISTTEGRIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 18.
DR PROSITE; PS00022; EGF_1; UNKNOWN_33.
DR PROSITE; PS01186; EGF_2; 25.
DR PROSITE; PS01187; EGF_CA; 16.
FT NON TER 1
SQ SEQUENCE 2281 AA; 239472 MW; DE23B2F129709D75 CRC64;

Query Match 34.88; Score 126; DB 4; Length 2281;
Best Local Similarity 59.08; Pred. No. 1.2e-06;
Matches 23; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

OY 26 CDPNCPENGICLPLGLAVGSCFCECPDGFDPNCSSVVE 64
DB 813 CDPNCPCLNGSCQDQ--VGFSFCSCLPFGAGPRCARDVD 849
ID 0916L8 PRELIMINARY; PRT; 2321 AA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gunel M., Aravanis-Tsakonas S.;
RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF058890; AAC14346.1;
DR EMBL; AF058881; AAC14346.1; JOINED.
DR EMBL; AF058882; AAC14346.1; JOINED.
DR EMBL; AF058883; AAC14346.1; JOINED.
DR EMBL; AF058884; AAC14346.1; JOINED.
DR EMBL; AF058885; AAC14346.1; JOINED.
DR EMBL; AF058886; AAC14346.1; JOINED.
DR EMBL; AF058887; AAC14346.1; JOINED.
DR EMBL; AF058888; AAC14346.1; JOINED.
DR EMBL; AF058889; AAC14346.1; JOINED.
DR EMBL; AF058890; AAC14346.1; JOINED.
DR EMBL; AF058891; AAC14346.1; JOINED.
DR EMBL; AF058892; AAC14346.1; JOINED.
DR EMBL; AF058893; AAC14346.1; JOINED.
DR EMBL; AF058894; AAC14346.1; JOINED.
DR EMBL; AF058895; AAC14346.1; JOINED.
DR EMBL; AF058896; AAC14346.1; JOINED.
DR EMBL; AF058897; AAC14346.1; JOINED.
DR EMBL; AF058898; AAC14346.1; JOINED.
DR EMBL; AF058899; AAC14346.1; JOINED.
DR HSSP; P00740; 1IXA.

```

DR INTERPRO: IPR000152; -
DR INTERPRO: IPR000561; -
DR INTERPRO: IPR000742; -
DR INTERPRO: IPR000800; -
DR INTERPRO: IPR001438; -
DR INTERPRO: IPR001881; -
DR INTERPRO: IPR002049; -
DR INTERPRO: IPR002110; -
DR PFAM: PF000008; EGF: 34.
DR PFAM: PF000023; ank: 6.
DR PFAM: PF00066; notch: 3.
DR PRINTS: PR00010; EGFLOOD.
DR PRINTS: PR00011; EGFPLAININ.
DR PROSITE: PS000010; ASK_HIDROXYL; 18.
DR PROSITE: PS000022; EGF_1; UNKNOWN_33.
DR PROSITE: PS01186; EGF_2; 25.
DR PROSITE: PS01187; EGF_CA; 16.
KW Glycoprotein; EGF-like domain.
SEQUENCE 2321 AA; 243629 MW; 3E70EC12A59CD638 CRC64

	Query Match	34.8%	Score 126;	DB 4;	Length 2321;
	Best Local Similarity	59.0%	Pred. No. 1.2e-06;		
Matches	23; Conservative	2;	Mismatches	12;	Indels 2; Gaps 1
OY	26 CDPNPCENGGICLPGLAVGSFSCDEPDTIDPNCSAYVE	64			
Db	853 CDPNPCLNCGSCDDG--VGSFSCICLCPGAGACRCARDVD	889			

RESULT	6			
090M47				
ID	090M47	PRELIMINARY;	PRT:	2321 AA.
AC	Q9UM47			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	NOTCH3.			
GN	NOTCH3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98049753; PubMed=8878478;			
RA	Aloul A., Corpechot C., Ducros A., Vahedi K., Chabriat H., Mouton P.,			
RA	Alamowitch S., Domenga V., Cecillion M., Marechal E., Maciazek J.,			
RA	Vaysiere C., Crnaud C., Cabanis E.A., Ruchoux M.M., Weissenbach J.,			
RA	Bach J.F., Bousser M.G., Tournier-Lasserre E.;			
RT	"Notch3 mutations in CADASIL, a hereditary adult-onset condition			
RT	causing stroke and dementia.";			
RL	Nature 383:707-710(1996).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=98049753; PubMed=9388399;			
RA	Joutel A., Vahedi K., Corpechot C., Trosch A., Chabriat H.,			
RA	Vaysiere C., Crnaud C., Maciazek J., Weissenbach J., Bousser M.G.,			
RA	Bach J.F., Tournier-Lasserre E.;			
RT	"Strong clustering and stereotyped nature of Notch3 mutations in			
RT	CADASIL patients.";			
RL	Lancet 350:1511-1515(1997).			
RL	EMBL: U97669; AAB91371.1; -.			
DR	HSSP: P00740; IEDM.			
DR	INTERPRO: IPR000152; -.			
DR	INTERPRO: IPR000561; -.			
DR	INTERPRO: IPR000742; -.			
DR	INTERPRO: IPR000800; -.			
DR	INTERPRO: IPR001438; -.			
DR	INTERPRO: IPR001881; -.			
DR	INTERPRO: IPR002049; -.			
DR	INTERPRO: IPR002110; -.			
DR	PFAM: PF00008; EGF; 34.			

DR PFAM: pf00023; ank: 6.
DR PFAM: pf00066; notch: 3.
DR PRINTS: PR00010; EGFLAOD.
DR PRINTS: PR00011; EGFAMITIN
DR PROSITE: PS00010; ASX_HYDROXYL_18.
DR PROSITE: PS00022; EGF-1_UNKNONN_33.
DR PROSITE: PS01186; EGF-2_25.
DR PROSITE: PS01187; EGF_1CA_15.
DR SEQUENCE 2321 AA; A4BPF2ECABTC836 CRC64

	Query March	34.8%;	Score 126;	DB 4;	Length 2321;
	Best Local Similarity	59.0%;	Pred. No. 1.2e-06;		
	Matches 22;	Conservative	2;	Mismatches 12;	Indels 2;
					Gaps 1
Qy	26	CDPNCENGICILPLGLAVGSFSCGCDGTTDNCSSVYE	64		
Db	853	CDPNCPLNGGSCDDG--VGSFSCSCILPGAGRCARDVD	889		

RESULT	7		
ID	Q25059	PRELIMINARY:	PRT; 406 AA.
AC	O25059;		
DT	01-NOV-1996 (TREMBLrel_01)	(Created)	
PT	01-NOV-1996 (TREMBLrel_01)	Last sequence update)	
DE	F1BROPELLIN III (FRAGMENT). 15.	Last annotation update)	
OS	Helicoidalis erythrogramma (Sea urchin).		
OC	Echinozoa; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;		
CC	Echinoidea; Euechinozoa; Echinacea; Echinolida; Echinometridae;		
OX	NCHI_taxid=7634;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RL	Bisgrove B.W.;		
DR	Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: L33862; AAA29996.1; -.		
DR	HSSP: P00740; 1EDM.		
DR	INTERPRO: IPR000083; -.		
DR	INTERPRO: IPR000088; -.		
DR	INTERPRO: IPR000152; -.		
DR	INTERPRO: IPR000561; -.		
DR	INTERPRO: IPR000742; -.		
DR	INTERPRO: IPR001010; -.		
DR	INTERPRO: IPR001438; -.		
DR	INTERPRO: IPR001861; -.		
DR	INTERPRO: IPR001947; -.		
DR	INTERPRO: IPR02049; -.		
DR	PFAM: PF00008; EGF_7.		
DR	PFAM: PF01382; Avidin; 1.		
DR	PRINTS: PRO0010; EGRBLOOD.		
DR	PRINTS: PRO0011; EGFLAMININ.		
DR	PRINTS: PRO0012; ENTPEPI.		
DR	PRINTS: PRO0286; CHARYBDOTOXIN.		
DR	PRINTS: PRO0287; THIOININ.		
DR	PRINTS: PRO0709; AVIDIN.		
DR	PROSITE: PS00010; ASK_HYDROXYL; UNKNOWN_7.		
DR	PROSITE: PS00022; EGF_1; UNKNOWN_7.		
DR	PROSITE: PS00577; AVIDIN; 1.		
DR	PROSITE: PS01186; EGF_2; 6.		
DR	PROSITE: PS01187; EGF_Ca; 5.		
KW	Glycoprotein; EGF-like domain.		
FT	NON_TER 1		
SQ	SEQUENCE 406 AA; 43475 MW EC98CCEBAB7E89A CRC64;		

```

Query Match:      34.3%; Score 124; DB 5; Length 406;
Best Local Similarity 46.5%; Pred. No. 3.8e-07;
Matches 20; Conservative 7; Mismatches 14; Indels 2; Gaps 1.

OY 23 GDICDPNCPENGICLPGLAVGSPSCPEPDGFTDPCNSVYEV 65
    | | | | | | | | | | | | | | | | | | | | | | | |

```

DB 13 GDDCNPNCQNGACID--QVNDYECIOPEFTGDCNCETDIDV 53

RESULT 8

Q9VB65 PRELIMINARY; PRT: 1404 AA.

AC Q9VB65;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE SER PROTEIN.

GN SER.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxId=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RC MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.C., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abrell J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,

RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,

RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Kethum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Moutklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacible J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

RL EMBL: AEO03759; AAF56678.1; -.

DR HSSP: P00743; 1CCF.

DR FLYBASE: FBgn0004197; Ser.

DR INTERPRO: IPR000152; -.

DR INTERPRO: IPR000255; -.

DR INTERPRO: IPR000515; -.

DR INTERPRO: IPR000561; -.

DR INTERPRO: IPR000742; -.

DR INTERPRO: IPR001438; -.

DR INTERPRO: IPR001774; -.

DR INTERPRO: IPR001881; -.

DR INTERPRO: IPR002049; -.

DR PRAM: PR00008; EGF_11.

DR PRAM: PF01414; DSL; 1.

DR PRINTS: PR00010; EGFBL00D.

DR PRINTS: PR00011; EGFPLAMIN.

DR PROSITE: PS00010; ASX_HYDROXYL; 7.

DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.

DR PROSITE: PS00022; EGF_1; 14.

DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.

DR PROSITE: PS01186; EGF_2; 8.

DR PROSITE: PS01187; EGF_CA; 5.

SQ SEQUENCE 1404 AA; 150342 MW; E988604001DAAAC84 CRC64;

Query Match 33.4%; Score 121; DB 5; Length 1404;

Best Local Similarity 60.0%; Pred. No. 3e-06;

Matches 21; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

QY 26 CDPNPGNGICLPLGLAVGSCPCPDFTPNCSYVE 60

DB 799 CSPNPGNGICLID--DGDFTCECMGWTGKRC 831

RESULT 9

Q9U0E2 PRELIMINARY; PRT: 585 AA.

AC Q9U0E2;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE SIGNAL RECEPTOR PROTEIN (FRAGMENT).

GN NOTCH.

OS Tribolium castaneum (Red flour beetle).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;

OC Cucujiformia; Tenebrionidae; Tribolium.

OX NCBI_TaxId=7070;

RN [1]

RP SEQUENCE FROM N.A.

RA Tautz D., Lardelli M., Westin J., Tamme R.,

RT "Embryonic expression of Tribolium Notch."

RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AJ005083; CAB65469.1; -.

DR HSSP: P00740; 1EDM.

DR INTERPRO: IPR000083; -.

DR INTERPRO: IPR000152; -.

DR INTERPRO: IPR000561; -.

DR INTERPRO: IPR000742; -.

DR INTERPRO: IPR000800; -.

DR INTERPRO: IPR001010; -.

DR INTERPRO: IPR001438; -.

DR INTERPRO: IPR001881; -.

DR INTERPRO: IPR002049; -.

DR PRAM: PF00008; EGF_11.

DR PRAM: PF00066; notch; 2.

DR PRINTS: PR00010; EGFBL00D.

DR PRINTS: PR00012; EGFPLAMIN.

DR PRINTS: PR00013; ENTPEI.

DR PRINTS: PR00287; THIONIN.

DR PROSITE: PS00010; ASX_HYDROXYL; 6.

DR PROSITE: PS00022; EGF_1; UNKNOWN_12.

DR PROSITE: PS01186; EGF_2; 9.

DR PROSITE: PS01187; EGF_CA; 4.

KW Receptor.

FT NON_TER 1 1

FT NON_TER 585 585

SQ SEQUENCE 585 AA; 63337 MW; 4CF7A51D0820D048 CRC64;

Query Match 32.3%; Score 117; DB 5; Length 585;

Best Local Similarity 51.3%; Pred. No. 3.8e-06;

Matches 20; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

QY 26 CDPNPGNGICLPLGLAVGSCPCPDFTPNCSYVE 64

DB 188 CDSAPCONGSTCID--LIGSYCKCKPKGFGONCCLNVD 224

```

RESULT 10
Q25058
ID Q25058 PRELIMINARY; PRT; 529 AA.
AC Q25058;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE FIBROPELIN IA (FRAGMENT).
OS Helicoidaris erythrograna (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinzoa;
OC Echinoidea; Echinoidae; Echinacea; Echinoida; Echinometridae;
OC Helicoidaris.
OX NCBI_TaxID=7634;
RN
RP SEQUENCE FROM N.A.
RA Bisgrove B.W.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: I33861; AAA29995.1; -.
DR HSSP: P08709; 1BP9.
DR INTERPRO: IPRO00083; -.
DR INTERPRO: IPRO00088; -.
DR INTERPRO: IPRO00152; -.
DR INTERPRO: IPRO00561; -.
DR INTERPRO: IPRO00742; -.
DR INTERPRO: IPRO01010; -.
DR INTERPRO: IPRO01438; -.
DR INTERPRO: IPRO01881; -.
DR INTERPRO: IPRO01947; -.
DR INTERPRO: IPRO02049; -.
DR PFAM: PF00008; EGF_10.
DR PFAM: PF01382; Avidin; 1.
DR PRINTS: PRO0010; EGFBLD.
DR PRINTS: PRO0011; EGFLAMININ.
DR PRINTS: PRO0012; ENTYPEI.
DR PRINTS: PRO0286; CHARYBTOXIN.
DR PRINTS: PRO0287; THIONIN.
DR PRINTS: PRO0709; AVIDIN.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_8.
DR PROSITE: PS00022; EGF_1; UNKNOWN_11.
DR PROSITE: PS00577; AVIDIN; 1.
DR PROSITE: PS01186; EGF_2; 10.
DR PROSITE: PS01187; EGF_CA; 7.
KM Glycoprotein; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 529 AA; 55543 MW; DAAE958FCF9ACB5A CRC64;

Query Match 31.5%; Score 114; DB 5; Length 529;
Best Local Similarity 43.9%; Pred. No. 7.9e-06;
Matches 18; Conservative 8; Mismatches 13; Indels 2; Gaps 1;

QY 24 DICDPNCGNGICLPLGLAVGSCFCEPDGFTDNCSSVYE 64
DB 175 DECAPSPCENGSGICLPG--VNGITCNCSGYTGNCETED 213

RESULT 11
Q13149
ID Q13149 PRELIMINARY; PRT; 2447 AA.
AC Q13149;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE NOTCH 2 (FRAGMENT).
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphae; Acanthopterygii; Petcomorphia; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN

```

```

RP SEQUENCE FROM N.A.
RA Nakamura T., Trowsdale J.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB004829; BAA20535.1; -.
DR HSSP: P00740; 1EDM.
DR INTERPRO: IPRO00083; -.
DR INTERPRO: IPRO00152; -.
DR INTERPRO: IPRO00561; -.
DR INTERPRO: IPRO00742; -.
DR INTERPRO: IPRO00800; -.
DR INTERPRO: IPRO01010; -.
DR INTERPRO: IPRO01438; -.
DR INTERPRO: IPRO01881; -.
DR INTERPRO: IPRO01947; -.
DR INTERPRO: IPRO02049; -.
DR INTERPRO: IPRO02110; -.
DR PFAM: PF00008; EGF_35.
DR PFAM: PF00023; ank; 6.
DR PFAM: PF00066; notch; 3.
DR PRINTS: PRO0010; EGFBLD.
DR PRINTS: PRO0011; EGFLAMININ.
DR PRINTS: PRO0012; ENTYPEI.
DR PRINTS: PRO0286; CHARYBTOXIN.
DR PRINTS: PRO0287; THIONIN.
DR PROSITE: PS00010; ASX_HYDROXYL; 22.
DR PROSITE: PS00022; EGF_1; UNKNOWN_35.
DR PROSITE: PS01186; EGF_2; 29.
DR PROSITE: PS01187; EGF_CA; 20.
KM Glycoprotein; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 2447 AA; 262542 MW; 1A8E2A372A085D84 CRC64;

Query Match 31.5%; Score 114; DB 13; Length 2447;
Best Local Similarity 58.3%; Pred. No. 3.6e-05;
Matches 21; Conservative 2; Mismatches 11; Indels 2; Gaps 1;

QY 24 DICDPNCGNGICLPLGLAVGSCFCEPDGFTDNC 59
DB 867 DDCSPNCLNGSGCYDD--VGSFSCFCEPGEHC 900

RESULT 12
Q70534
ID Q70534 PRELIMINARY; PRT; 383 AA.
AC Q70534;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE ZOG.
DE ZOG.
GN ZOG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN
RP SEQUENCE FROM N.A.
RA STRAIN-SPRAGUE-DAWLEY;
RA Halder S.K., Takemori H., Akira W., Yasuki N., Mitsuhiro O.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: D84336; BAA25881.1; -.
DR HSSP: P00740; 1IXA.
DR INTERPRO: IPRO00152; -.
DR INTERPRO: IPRO00561; -.
DR INTERPRO: IPRO00742; -.
DR INTERPRO: IPRO01881; -.
DR PFAM: PF00008; EGF_5.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_5.
DR PROSITE: PS01186; EGF_2; 6.
SQ SEQUENCE 383 AA; 41149 MW; 123724C9801A9FD5 CRC64;

```


4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: May 23, 2001, 06:15:14 ; Search time 75.05 Seconds
(without alignments)
49,508 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Perfect score: 362
Sequence: 1 MKRSVAWLVGLSLGVPQF.....FSCGCPDGTDPNCSSVVEV 65

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_0401:*

- 1: /SID56/gcgcdata/geneseq/geneseqp/AA1980.DAT:*
- 2: /SID56/gcgcdata/geneseq/geneseqp/AA1981.DAT:*
- 3: /SID56/gcgcdata/geneseq/geneseqp/AA1982.DAT:*
- 4: /SID56/gcgcdata/geneseq/geneseqp/AA1983.DAT:*
- 5: /SID56/gcgcdata/geneseq/geneseqp/AA1984.DAT:*
- 6: /SID56/gcgcdata/geneseq/geneseqp/AA1985.DAT:*
- 7: /SID56/gcgcdata/geneseq/geneseqp/AA1986.DAT:*
- 8: /SID56/gcgcdata/geneseq/geneseqp/AA1987.DAT:*
- 9: /SID56/gcgcdata/geneseq/geneseqp/AA1988.DAT:*
- 10: /SID56/gcgcdata/geneseq/geneseqp/AA1989.DAT:*
- 11: /SID56/gcgcdata/geneseq/geneseqp/AA1990.DAT:*
- 12: /SID56/gcgcdata/geneseq/geneseqp/AA1991.DAT:*
- 13: /SID56/gcgcdata/geneseq/geneseqp/AA1992.DAT:*
- 14: /SID56/gcgcdata/geneseq/geneseqp/AA1993.DAT:*
- 15: /SID56/gcgcdata/geneseq/geneseqp/AA1994.DAT:*
- 16: /SID56/gcgcdata/geneseq/geneseqp/AA1995.DAT:*
- 17: /SID56/gcgcdata/geneseq/geneseqp/AA1996.DAT:*
- 18: /SID56/gcgcdata/geneseq/geneseqp/AA1997.DAT:*
- 19: /SID56/gcgcdata/geneseq/geneseqp/AA1998.DAT:*
- 20: /SID56/gcgcdata/geneseq/geneseqp/AA1999.DAT:*
- 21: /SID56/gcgcdata/geneseq/geneseqp/AA2000.DAT:*
- 22: /SID56/gcgcdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	362	100.0	481	18	W10365	Human developmenta
2	362	100.0	513	20	W94685	Human Del-1 protei
3	304	84.0	221	18	W10366	Murine Del-1 trunc
4	304	84.0	221	20	W94684	Truncated murine D
5	304	84.0	480	18	W10364	Mouse developmenta
6	304	84.0	480	20	W94683	Murine Del-1 prote
7	245	67.7	102	18	W10370	Human Del-1 splice
8	222.5	61.5	57	20	W94686	Human Del-1 epiderm
9	127	35.1	25	18	W10368	Human Del-1 signal
10	127	35.1	25	20	W94698	Human Del-1 signal
11	126	34.8	2321	19	W49698	Human Notch3 prote

12	121	33.4	1404	21	Y59600	Drosophila Serrate
13	113	31.2	385	15	R56167	Neuroendocrine tum
14	113	31.2	512	20	W75494	Truncated human de
15	113	31.2	520	18	W18348	Proliferation and
16	113	31.2	702	18	W18349	Proliferation and
17	113	31.2	702	20	W75495	Truncated human de
18	113	31.2	723	18	W18353	Proliferation and
19	113	31.2	723	20	W94498	Human delta-1 prot
20	113	31.2	723	20	W75492	Human delta-1 prot
21	113	31.2	723	21	B00172	Human delta-1 prote
22	113	31.2	723	21	B33482	Human PRO172 prote
23	113	31.2	723	21	B24388	Human PRO172 prote
24	113	31.2	723	21	B83227	PRO172 polypeptide
25	113	31.2	723	21	V79032	Human delta protei
26	113	31.2	723	22	B53064	Human angiogenesis
27	112	30.9	46	17	R96227	Novel growth facto
28	112	30.9	383	15	R56166	Neuroendocrine tum
29	111	30.7	157	18	W11730	H-Delta-1 polypept
30	111	30.7	660	18	W11725	H-Delta-1 polypept
31	111	30.7	722	21	V79028	Murine delta prote
32	111	30.7	1404	14	R38304	Sequence of a serr
33	109	30.1	722	18	W11720	M-Delta-1 polypept
34	109	30.1	1055	19	W44298	Human serrate 2 pr
35	109	30.1	1148	20	W87895	Human JAGGED2 prot
36	109	30.1	1212	19	W44299	Human serrate 2
37	109	30.1	1257	17	W05834	Human Serrate-2 (H
38	109	30.1	1257	21	V59598	Human Serrate prot
39	108	29.8	78	21	B41310	Human OREF ORF1074
40	108	29.8	149	21	B40789	Human OREF ORF553
41	106	29.3	46	17	R96226	Novel growth facto
42	106	29.3	412	16	R64266	Human Factor-IX.
43	106	29.3	694	21	B42900	Human OREF ORF2664
44	105.5	29.1	415	20	V03203	Amino acid sequenc
45	105.5	29.1	454	16	R67710	Human Factor-IX.

ALIGNMENTS

RESULT 1	ID	W10365	standard: Protein: 481 AA.
XX	XX	W10365:	
DT	DT	03-MAY-1997 (first entry)	
XX	XX		Human developmentally-regulated endothelial cell locus-1 protein.
DE	DE		
XX	XX		Del-1; developmentally-regulated endothelial cell locus-1;
KW	KW		signal transduction; cancer; tumour marker; angiogenesis;
KW	KW		diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW	KW		cardiac ischemia; stroke; vascular disease; wound healing;
KW	KW		vulnerable; bone formation; diagnosis; therapy.
XX	XX		Homo sapiens.
OS	OS		
XX	XX		
FH	FH		Location/Qualifiers
FT	FT	Key	3..23
FT	FT	Peptide	/label= Sig_peptide
FT	FT	Protein	24..481
FT	FT		/label= Mat_protein
FT	FT	Domain	26..59
FT	FT		/label= EGF-1
FT	FT		/note= "epidermal growth factor-like domain 1"
FT	FT	Domain	78..116
FT	FT		/label= EGF-2
FT	FT		/note= "epidermal growth factor-like domain 2"
FT	FT	Domain	123..154
FT	FT		/label= EGF-3
FT	FT		/note= "epidermal growth factor-like domain 3"
FT	FT	Domain	158..315
FT	FT		/label= Discoidin-1

```

FT      /note="Discoidin I/factor VIII-like domain 1"
FT      320..477
FT      Domain
FT      /label="Discoidin-2"
FT      /note="Discoidin I/factor VIII-like domain 2"

XX      W09640769-A1.
XX
XX      19-DEC-1996.
XX
XX      05-JUN-1996; 96WO-US09456.
XX
XX      07-JUN-1995; 95US-0480229.
XX
XX      (PROG-) PROGENITOR INC.
XX      (UYVA-) UNIV VANDERBILT.
XX
XX      Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
XX      WPI; 1997-052233/05.
XX      N-PSDB; T47343.
XX
XX      New developmentally regulated endothelial cell locus-1 (del-1) gene
XX      - used to develop prods. for the diagnosis and treatment of cancer
XX      and conditions involving abnormal angiogenesis
XX
XX      Claim 4; Fig 6; 137pp: English.
XX
XX      Human Del-1 (W10365) is the polypeptide product of the human
XX      developmentally-regulated endothelial cell locus-1 (del-1) gene
XX      (T47343). It shows 94% amino acid homology to the mouse Del-1
XX      protein (W10364). Structurally, members of this novel gene family
XX      contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
XX      domains. Del-1 is expressed in endothelial and certain tumour
XX      cells. Its ability to inhibit vascular formation allows its used
XX      as an anti-angiogenic agent. It can be used as a tumour marker,
XX      to identify Del-1 binding partners, and to modulate endothelial
XX      cell growth and blood vessel formation. Recombinant Del-1 can be
XX      produced in transformed host cells utilizing vectors incorporating
XX      del-1 nucleic acids.
XX
XX      Sequence 481 AA:
XX
SQ      Query Match 100.0%; Score 362; DB 18; Length 481;
SQ      Best Local Similarity 100.0%; Pred. No. 5.8e-25;
SQ      Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
SQ
SQ      QY 1 MKRSVAWLVIGSLGVPOFGKIDCDPMPGNGSICLPGLAVSFSECPDGTDPNCS 60
SQ      I |||||
SQ      I MKRSVAWLVIGSLGVPGTGYK9dcdpmpcngsngiclpglavsgfsccepdgtdpncs 60
SQ
SQ      QY 61 SVVEV 65
SQ      |||||
SQ      Db 61 svvev 65
SQ
SQ      RESULT 2
SQ      W94685
SQ      ID W94685 standard; Protein; 513 AA.
SQ
SQ      AC W94685;
SQ
SQ      DT 05-MAY-1999 (first entry)
SQ
SQ      DE Human Del-1 protein.
SQ
SQ      KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
SQ      discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
SQ      diabetic retinopathy; rheumatoid arthritis; endometriosis;
SQ      angiogenesis.
SQ
SQ      KW Homo sapiens.
SQ

```

[illegible]

[illegible]

XX discoidin I; factor VIII-like domain; epidermal growth factor; bFGF;
KM diabetic retinopathy; rheumatoid arthritis; endometriosis;
KM angiogenesis.
OS Mus sp.
XX
XX US5877281.N.
XX PD
XX 02-MAR-1999.
XX PF 05-JUN-1996; 96US-0659235.
XX XX 05-JUN-1996; 96US-0659235.
PR 07-JUN-1995; 95US-0480229.
XX (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B., Quettermous T., Snodgrass HR., Zupancic TJ;
DR WFI: 1999-189720/16.
DR N-PDB: X18507.
XX
PF Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
PS Claim 3: Column 71-72; 73pp; English.

CC The present sequence is truncated murine developmentally-regulated
CC endothelial cell locus 1 (Del-1). The protein has epidermal growth factor
CC (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1
CC proteins have an inhibitory effect on angiogenesis (blood vessel growth)
CC this activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha v beta 3
CC integrin, and is an apoptosis inducer.

XX Sequence 221 AA:
SQ

Query Match 84.0%; Score 304; DB 20; Length 221;
Best local similarity 84.6%; Pred. No. 4.1e-20;
Matches 55; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Oy 1 MKRSVAVMLVLGLSLGVDFPGKGDIDPPPCENGICLPGLIAGVSFCBPDPGFDPNCS 60
||| |||||||||||||||||||||||||||| | ||||| :||| |||
Db 1 mkhlvaewllvglslypgfgkgdlnpcncenglclsgldadsfcscpcgfaqgnps 60

Oy 61 SYVEV 65
|||||
Db 61 SYVEV 65

DE Mouse developmentally-regulated endothelial cell locus-1 protein.
XX
XX
RM Del-1; developmentally-regulated endothelial cell locus-1;
KM signal transduction; cancer; tumor marker; angiogenesis;
KM diabetetic retinopathy; rheumatoid arthritis; endometriosis;
KM cardiac ischaemia; stroke; vascular disease; wound healing;
KM vulnery; bone formation; diagnosis; therapy.

```

XX Mus sp.
OS
XX
XX Key location/Qualifiers
XX Peptide 1..23
XX 24..480
XX Protein /label= sig-Deptide
XX /label= Mat-protein
XX Domain 26..39
XX /label= EGF-1
XX /note= "epidermal growth factor-like domain 1"
XX Domain 78..116
XX /label= EGF-2
XX /note= "epidermal growth factor-like domain 2"
XX Domain 123..154
XX /label= EGF-3
XX /note= "epidermal growth factor-like domain 3"
XX Domain 158..314
XX /label= discoidin-1
XX /note= "discoidin I/factor VIII-like domain 1"
XX Domain 319..476
XX /label= discoidin-2
XX /note= "discoidin I/factor VIII-like domain 2"
XX
XX WC0640769-A1.
XX
XX 19-DEC-1996.
XX
XX 05-JUN-1996; 96WO-US09456.
XX
XX 07-JUN-1995; 95US-0480229..
XX
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quatermous T, Snodgrass HR, Zupancic TJ;
XX
XX WPI: 1997-052233/05.
XX N-PSDB: T47338.
XX
XX New developmentally regulated endothelial cell locus-1 (del-1) gene
XX - used to develop prods. for the diagnosis and treatment of cancer
XX and conditions involving abnormal angiogenesis
XX
XX Claim 3; Fig 6; 137pp: English.
XX
XX Murine Del-1 (W10364) is the polypeptide product of the murine
XX developmentally-regulated endothelial cell locus-1 (del-1) gene
XX (747338). It shows 94% amino acid homology to the human Del-1
XX protein (W10365). Structurally, members of this novel gene family
XX contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
XX domains. Del-1 is expressed in endothelial and certain tumour
XX cells. Its ability to inhibit vascular formation allows its used
XX as an anti-angiogenic agent. It can be used as a tumour marker,
XX to identify Del-1 binding partners, and to modulate endothelial
XX cell growth and blood vessel formation. Recombinant Del-1 can be
XX produced in transformed host cells utilizing vectors incorporating
XX del-1 nucleic acids.
XX
XX Sequence 480 AA:
SQ

```

Query Match 84.0%; Score 304; DB 18; Length 480;
 Best Local Similarity 84.6%; Pred. No. 8.2e-20;
 Matches 55; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

OY 1 MKRSVAWMLVGLSLGVPQFGKGDICDPNCPENGIGICLPGLAVGSPSCPDGFTPNCS 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 mkhlvaawllvgslslypqfgykdicnppcengiglsjladstfscpcpfgapncs 60
OY 61 SYVEV 65
   |||||
Db 61 syvev 65

```

```

RESULT 6
W94683
XX W94683 standard: Protein; 480 AA.
XX
XX W94683:
XX
XX 05-MAY-1999 (first entry)
XX
XX Murine Del-1 protein.
XX
XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
XX diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX angiogenesis.
XX
XX Mus sp.
XX
XX US877281-A.
XX
XX 02-MAR-1999.
XX
XX 05-JUN-1996; 96US-0659235.
XX
XX 05-JUN-1996; 96US-0659235.
XX
XX 07-JUN-1995; 95US-0480229.
XX
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Quatermous T, Snodgrass HR, Zupancic TJ;
XX
XX WPI: 1999-189720/16.
XX N-PSDB: X18506.
XX
XX Del-1 polypeptide sequences - useful for treatment of cancer,
XX diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX Claim 1; Fig 3; 73pp: English.
XX
XX The present sequence is murine developmentally-regulated endothelial
XX cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like
XX domains and discoidin I/factor VIII-like domains. The Del-1 proteins
XX have an inhibitory effect on angiogenesis (blood vessel growth), this
XX activity may be useful clinically to prevent neovascularisation of
XX tissues such as tumour nodules and prevention of metastases. The anti-
XX angiogenic activity of Del-1 may be used to treat abnormal conditions
XX that result from angiogenesis, including cancer, diabetic retinopathy,
XX rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX stroke, wound healing and peripheral vascular disease. Del-1 is also
XX useful for promoting bone formation. Del-1 binds to alpha V beta 3
XX integrin, and is an apoptosis inducer.
XX
XX Sequence 480 AA:
SQ

```

Query Match 84.0%; Score 304; DB 20; Length 480;
 Best Local Similarity 84.6%; Pred. No. 8.2e-20;
 Matches 55; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

OY 1 MKRSVAWMLVGLSLGVPQFGKGDICDPNCPENGIGICLPGLAVGSPSCPDGFTPNCS 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 mkhlvaawllvgslslypqfgykdicnppcengiglsjladstfscpcpfgapncs 60
OY 61 SYVEV 65
   |||||
Db 61 syvev 65

```

RESULT 7
 W10370

```

ID W10370 standard; Protein; 102 AA.
XX
AC W10370;
XX
DT 03-MAY-1997 (first entry)
XX
DE Human Del-1 splice variant partial sequence.
XX
KW Del-1; developmentally-regulated endothelial cell locus-1;
KW signal transduction; cancer; tumour marker; angiogenesis;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW cardiac ischaemia; stroke; vascular disease; wound healing;
KW vulnereary; bone formation; diagnosis; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH 3..36
FH Domain /label= "epidermal growth factor-like domain 1"
FH /note= "epidermal growth factor-like domain 1"
FH 45..83
FH Domain /label= "epidermal growth factor-like domain 2"
FH /note= "epidermal growth factor-like domain 2"
FH 90..102
FH Domain /label= "EGF-3"
FH /note= "N-terminal portion of epidermal growth
FH factor-like domain 3"
FH
XX
XX W09640769-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96MO-US09456.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
XX (DYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1997-052233/05.
XX
DR New developmentally regulated endothelial cell locus-1 (del-1) gene
PT - used to develop prods. for the diagnosis and treatment of cancer
PT and conditions involving abnormal angiogenesis
PT
XX
PS Example; Page 91; 137pp; English.
XX
CC A polypeptide sequence (W10370) shows a portion of a variant form
CC of human developmentally-regulated endothelial cell locus-1 (del-1)
CC polypeptide. In comparison with the major form (W10365) of human
CC Del-1, 10 amino acid residues are missing between EGF-like domains
CC 1 and 2 of the Del-1 variant. This is a result of alternative
CC splicing (see also T47340).
XX
SQ Sequence 102 AA:

```

```

Query Match 67.7%; Score 245; DB 18; Length 102;
Best Local Similarity 100.0%; Pred. No. 3.7e-15;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 24 DICDPCNGGICLPGILAVSFSCECPDGTDPNCSSVVEV 65
   |||
DB 1 dicdpcnggiclpgilavgsfscpcpdtgtdpncssvvev 42

```

```

RESULT 8
W94686
ID W94686 standard; peptide: 57 AA.
XX
AC W94686;

```

```

XX
DT 05-MAY-1999 (first entry)
XX
DE Del-1 epidermal growth factor like domain #1.
XX
KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW angiogenesis.
XX
OS Mus sp.
XX Synthetic.
XX
PN US5877281-A.
XX
PD 02-MAR-1999.
XX
PF 05-JUN-1996; 96US-0659235.
XX
PR 05-JUN-1996; 96US-0659235.
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (DYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX
DR Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
PT
XX
PS Claim 40; Fig 10; 73pp; English.
XX
CC The present sequence represents an epidermal growth factor like domain
CC from developmentally-regulated endothelial cell locus 1 (Del-1). Del-1
CC also has discoidin I/factor VIII-like domains. The Del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
SQ Sequence 57 AA:

```

```

Query Match 61.5%; Score 222.5; DB 20; Length 57;
Best Local Similarity 88.9%; Pred. No. 2.2e-13;
Matches 40; Conservative 0; Mismatches 5; Indels 5; Gaps 1;

```

```

QY 26 CDPNCEGNGICLPGILAVSFSCECPDGTDPNCSSVVEV 65
   |||
DB 1 cdpncegngiclpgilavgsfscpcpdtgtdpncssvvev 45

```

```

RESULT 9
W10368
ID W10368 standard; Peptide: 25 AA.
XX
AC W10368;
XX
DT 03-MAY-1997 (first entry)
XX
DE Human Del-1 signal peptide.
XX
KW Del-1; developmentally-regulated endothelial cell locus-1;
KW signal transduction; cancer; tumour marker; angiogenesis;
KW diabetic retinopathy; rheumatoid arthritis; endometriosis;

```

KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnerability; bone formation; diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 3-21
 FT Peptide /label-Sig-peptide
 XX
 PN W09640769-A1.
 XX
 PD 19-DEC-1996.
 XX
 PD 05-JUN-1996; 96WO-US09456.
 XX
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIT VANDERBILT.
 XX
 PI Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
 XX
 DR WPI; 1997-052233/05.
 XX
 PT New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis
 XX
 PS Example: Fig 9; 137pp; English.
 XX
 CC The predicted amino acid sequence (W10368) at the N-terminus
 CC of the human developmentally-regulated endothelial cell locus-1
 CC (del-1) gene product (W10365) shows characteristics common to
 CC signal peptides. The putative signal begins with a basic Arg
 CC residue and is followed by a stretch of 18 hydrophobic residues.
 CC The Chou and Fasman algorithm predicts that the putative signal
 CC sequence is followed by a protein turn structure, a feature
 CC commonly found after signal peptides. The Del-1 protein is
 CC secreted by expressing cells.
 XX
 SQ Sequence 25 AA;
 XX

Query Match 35.1%; Score 127; DB 18; Length 25;
 Best Local Similarity 100.0%; Pred. No. 3.2e-05;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVWLVLGSLGVPOFGKDI 25
 DB 1 mkrsyavwllvgislgyvpqfgkdi 25

RESULT 10
 W94698
 ID W94698 standard; peptide; 25 AA.
 XX
 AC W94698:
 XX
 DT 05-MAY-1999 (first entry)
 XX
 DE Human Del-1 signal protein.
 XX
 DE Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW angiogenesis.
 XX
 OS Homo sapiens.
 XX
 PN US9877281-A.
 XX
 PD 02-MAR-1999.
 XX

PF 05-JUN-1996; 96US-0659235.
 XX
 PR 05-JUN-1996; 96US-0659235.
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIT VANDERBILT.
 XX
 PI Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
 XX
 DR WPI; 1999-189720/16.
 XX
 PT Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Example: Column 63-65; 73pp; English.
 XX
 CC The present sequence is the human developmentally-regulated endothelial
 CC cell locus 1 (del-1) amino terminus peptide sequence. The Del-1 protein
 CC has epidermal growth factor (EGF) like domains and discoidin I/factor
 CC VIII-like domains. The Del-1 proteins have an inhibitory effect on
 CC angiogenesis (blood vessel growth), this activity may be useful
 CC clinically to prevent neovascularisation of tissues such as tumour
 CC nodules and prevention of metastases. The anti-angiogenic activity of
 CC Del-1 may be used to treat abnormal conditions that result from
 CC angiogenesis, including cancer, diabetic retinopathy, rheumatoid
 CC arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be
 CC used to treat cardiac ischaemia, thrombotic stroke, wound healing and
 CC peripheral vascular disease. Del-1 is also useful for promoting bone
 CC formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis
 CC inducer.
 XX
 SQ Sequence 25 AA;
 XX

Query Match 35.1%; Score 127; DB 20; Length 25;
 Best Local Similarity 100.0%; Pred. No. 3.2e-05;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVWLVLGSLGVPOFGKDI 25
 DB 1 mkrsyavwllvgislgyvpqfgkdi 25

RESULT 11
 W49698
 ID W49698 standard; Protein; 2321 AA.
 XX
 AC W49698:
 XX
 DT 21-DEC-1998 (first entry)
 XX
 DE Human Notch3 protein.
 XX
 DE Human; Notch3; transmembrane receptor; lateral inhibition; regulation;
 KW developmental cascade; neurogenic gene; mutant; neurological disorder;
 KW cerebral autosomal dominant arteriopathy; subcortical infarct; CADASIL;
 KW leukoencephalopathy; therapy.
 XX
 OS Homo sapiens.
 XX
 PN FR2751986-A1.
 XX
 PD 06-FEB-1998.
 XX
 PF 16-APR-1997; 97FR-0004680.
 XX
 PR 01-AUG-1996; 96FR-0009733.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Bach JF, Bousser MG, Jourel A, Tournier Lasserre E;
 XX

Wed May 23 06:29:20 2001

us-09-237-981-14_copy_33_97.rag

Page 9

Search completed: May 23, 2001, 06:15:15
Job time: 229 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:18:13 ; Search time: 179.67 Seconds

(without alignments)
58.193 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Perfect score: 362
Sequence: 1 MKRSVAWLVLTGLSGVQF.....FSCPCPDGPTDNCSSVVEV 65

Scoring table: Biosum62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending Patents AA Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	362	100.0	481	10 US-08-659-235B-29	Sequence 29, Appl
2	362	100.0	513	8 US-08-480-229B-14	Sequence 14, Appl
3	362	100.0	513	10 US-08-659-235B-14	Sequence 14, Appl
4	362	100.0	513	16 US-09-237-981-14	Sequence 14, Appl
5	304	84.0	221	16 US-09-237-981-29	Sequence 29, Appl
6	304	84.0	480	8 US-08-480-229B-10	Sequence 10, Appl
7	304	84.0	480	10 US-08-659-235B-10	Sequence 10, Appl
8	304	84.0	480	16 US-09-237-981-10	Sequence 10, Appl
9	238	65.7	185	23 US-60-230-433-1806	Sequence 1806, Ap
10	238	65.7	203	23 US-60-233-644-77	Sequence 77, Appl

11	222.5	61.5	57	8 US-08-480-229B-23	Sequence 23, Appl
12	222.5	61.5	57	10 US-08-659-235B-23	Sequence 23, Appl
13	222.5	61.5	57	16 US-09-237-981-23	Sequence 23, Appl
14	127	35.1	25	8 US-08-480-229B-22	Sequence 22, Appl
15	127	35.1	25	10 US-08-659-235B-22	Sequence 22, Appl
16	127	35.1	25	16 US-09-237-981-22	Sequence 22, Appl
17	126	34.8	196	23 US-60-230-445-1215	Sequence 1215, Ap
18	126	34.8	2004	23 US-60-207-583-595	Sequence 595, Ap
19	126	34.8	2321	16 US-09-230-652-2	Sequence 2, Appl
20	126	34.8	2321	16 US-09-287-845-5	Sequence 2, Appl
21	126	34.8	2321	20 US-09-612-226-2	Sequence 2, Appl
22	121	33.4	1404	15 US-07-808-458-2	Sequence 2, Appl
23	121	33.4	1404	15 US-09-195-524-2	Sequence 255, App
24	121	33.4	1404	23 US-60-171-625-255	Sequence 10433, A
25	121	33.4	1404	23 US-60-173-464-10433	Sequence 12824, A
26	121	33.4	1404	23 US-60-191-637-12824	Sequence 10088, A
27	121	33.4	1404	23 US-60-191-681-10088	Sequence 12829, A
28	121	33.4	1407	23 US-60-167-217-12829	Sequence 3161, Ap
29	117	32.3	66	23 US-60-170-373-3161	Sequence 4202, Ap
30	117	32.3	77	23 US-60-373-4302	Sequence 575, App
31	117	32.3	192	23 US-60-207-214-575	Sequence 540, App
32	117	32.3	192	23 US-60-213-846-540	Sequence 250, App
33	117	32.3	460	23 US-60-205-420-250	Sequence 82, Appl
34	117	32.3	494	23 US-60-200-591-82	Sequence 120, Appl
35	117	32.3	673	23 US-60-201-718-120	Sequence 119, App
36	117	32.3	807	23 US-60-205-420-248	Sequence 249, App
37	117	32.3	851	23 US-60-201-718-119	Sequence 118, App
38	117	32.3	887	23 US-60-205-420-248	Sequence 118, App
39	117	32.3	903	23 US-60-201-718-118	Sequence 389, App
40	117	32.3	969	23 US-60-205-423-389	Sequence 117, Appl
41	117	32.3	972	23 US-60-201-718-117	Sequence 2, Appl
42	117	32.3	2444	16 US-09-287-845-2	Sequence 2, Appl
43	117	32.3	2556	4 PCT-US99-15817-2	Sequence 20, Appl
44	117	32.3	2556	4 US-08-083-590-20	Sequence 2, Appl
45	117	32.3	2556	12 US-08-899-232-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-659-235B-29
Sequence 29, Application US/08659235B

GENERAL INFORMATION:
APPLICANT: Quentemus, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-29

Query Match 100.0%; Score 362; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 5.8e-31;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGLAVGFSCECPDGFDPNCS 60
DB 1 MKRSYAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGLAVGFSCECPDGFDPNCS 60
QY 61 SYVEV 65
DB 61 SYVEV 65

RESULT 2
US-08-480-229B-14
Sequence 14, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480.229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229B-14

Query Match 100.0%; Score 362; DB 8; Length 513;
Best Local Similarity 100.0%; Pred. No. 6.2e-31;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGLAVGFSCECPDGFDPNCS 60
DB 33 MKRSYAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGLAVGFSCECPDGFDPNCS 92
QY 61 SYVEV 65
DB 93 SYVEV 97

RESULT 3
US-08-659-235B-14
Sequence 14, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659.235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-14

Query Match 100.0%; Score 362; DB 10; Length 513;
Best Local Similarity 100.0%; Pred. No. 6.2e-31;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSYAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGLAVGFSCECPDGFDPNCS 60
DB 33 MKRSYAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGLAVGFSCECPDGFDPNCS 92
QY 61 SYVEV 65
DB 93 SYVEV 97

RESULT 4
US-09-237-981-14
Sequence 14, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid

APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2271
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 790-9090
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
FILING DATE:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-14

Query Match 100.0%; Score 362; DB 16; Length 513;
Best Local Similarity 100.0%; Pred. No. 6,2e-31;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKRSVAVMLVGLSGVPGFGKGDICDPNCPENGICLPLGLAVGFSCECPGFTDPNCS 60
DB 33 MKRSVAVMLVGLSGVPGFGKGDICDPNCPENGICLPLGLAVGFSCECPGFTDPNCS 92

OY 61 SVEEV 65
DB 93 SVEEV 97

RESULT 5
US-09-237-981-29
Sequence 29, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Queternous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
FILING DATE:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-09-237-981-29

Query Match 84.0%; Score 304; DB 16; Length 221;
Best Local Similarity 84.6%; Pred. No. 4,9e-25;
Matches 55; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 MKRSVAVMLVGLSGVPGFGKGDICDPNCPENGICLPLGLAVGFSCECPGFTDPNCS 60
DB 1 MKHVAAMLVGLSLGVPGFGKGDICDPNCPENGICLPLGLAVGFSCECPGFTDPNCS 60

OY 61 SVEEV 65
DB 61 SVEEV 65

RESULT 6
US-08-480-229B-10
Sequence 10, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Queternous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.


```
US-60-230-435-1806
; Sequence 1806, Application US/602330435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1806
; LENGTH: 185
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(185)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-230-435-1806
```

```
Query Match 65.7%; Score 238; DB 23; Length 185;
Best Local Similarity 97.6%; Pred. No. 5.4e-18;
Matches 41; Conservative 1; Indels 0; Gaps 0;
```

```
OY 24 DCDPNCENGICLPLGLAVGFSFCECPDGTDPNCSSVVEY 65
|||||
DB 2 DCDPNCENGICLPLGLADSFSECPDGTDPNCSSVVEY 43
```

```
RESULT 10
US-60-233-644-77
; Sequence 77, Application US/60233644
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
; TITLE OF INVENTION: HORMONE RECEPTOR PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO00824
; CURRENT APPLICATION NUMBER: US/60/233,644
; CURRENT FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 203
; TYPE: PRT
; ORGANISM: HUMAN
US-60-233-644-77
```

```
Query Match 65.7%; Score 238; DB 23; Length 203;
Best Local Similarity 97.6%; Pred. No. 6e-18;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 24 DCDPNCENGICLPLGLAVGFSFCECPDGTDPNCSSVVEY 65
|||||
DB 12 DCDPNCENGICLPLGLADSFSECPDGTDPNCSSVVEY 53
```

```
RESULT 11
US-08-480-229B-23
; Sequence 23, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELTX: 6614 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-229B-23
```

```
Query Match 61.5%; Score 222.5; DB 8; Length 57;
Best Local Similarity 88.9%; Pred. No. 7.9e-17;
Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

```
OY 26 CDPNCEGICLPLGLAVG----SFSECPDGTDPNCSSVVEY 65
|||||
DB 1 CDPNCEGICLPLGLAVGXXXXXSFSCECPDGTDPNCSSVVEY 45
```

```
RESULT 12
US-08-659-235B-23
; Sequence 23, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-639-235B-23

Query Match 61.5%; Score 222.5; DB 10; Length 57;
Best Local Similarity 88.9%; Pred. No. 7.9e-17;
Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 26 CDPNCEGICLPGLAVG-----SFSCECPDGFDPNCSVYEV 65
DB 1 CDPNCEGICLPGLAVGXXXXXSFSCECPDGFDPNCSVYEV 45

RESULT 13

US-09-237-981-23
Sequence 23, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Polissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 09/09/237,981
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Polissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-237-981-23

Query Match 61.5%; Score 222.5; DB 16; Length 57;
Best Local Similarity 88.9%; Pred. No. 7.9e-17;

Matches 40; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
QY 26 CDPNCEGICLPGLAVG-----SFSCECPDGFDPNCSVYEV 65
DB 1 CDPNCEGICLPGLAVGXXXXXSFSCECPDGFDPNCSVYEV 45

RESULT 14

US-08-480-229B-22
Sequence 22, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Polissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Polissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:

Query Match 35.1%; Score 127; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.1e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAVWLVLVGLSLGVPQFGKDI 25
DB 1 MKRSVAVWLVLVGLSLGVPQFGKDI 25

RESULT 15

US-08-659-235B-22
Sequence 22, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-659-235B-22

Query Match 35.1%; Score 127; DB 10; Length 25;
Best local Similarity 100.0%; Pred. No. 7.1e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAVMLLVGSLGVPQFGKDI 25
|||||
Db 1 MKRSVAVMLLVGSLGVPQFGKDI 25

Search completed: May 23, 2001, 06:18:14
Job time: 408 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:54 ; Search time 28 Seconds

(without alignments)
16.640 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97

Perfect score: 362
Sequence: 1 MKRSVAWMLVGLSLGVPOF.....FSCBPDGFTDPCSSVVEV 65

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41833 seqs, 7168127 residues

Total number of hits satisfying chosen parameters: 41833

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	113	31.2	723	5	US-09-828-366-21
2	109	30.1	42	5	US-09-249-697A-7
3	105.5	29.1	461	1	PCT-US01-12012-3
4	102.5	28.3	1964	5	US-09-467-997-1
5	100	27.6	4044	6	US-60-248-505-791
6	98	27.1	310	1	PCT-US01-01302-103
7	95.5	26.4	110	1	PCT-US01-01302-89
8	95.5	26.4	122	1	PCT-US01-01302-97
9	95	26.2	188	5	PCT-US01-01302-109
10	92	25.4	440	5	US-09-467-997-2
11	91	25.1	383	1	PCT-US01-11988-770
12	91	25.1	383	5	US-09-833-245-770
13	90	24.9	226	1	PCT-US01-01302-92
14	88	24.3	48	5	US-09-802-704-11
15	88	24.3	415	5	US-09-802-704-7
16	88	24.3	439	1	PCT-US01-11988-627
17	88	24.3	439	1	PCT-US01-11988-627
18	88	24.3	439	5	US-09-802-704-4
19	88	24.3	439	5	US-09-833-245-627
20	88	24.3	439	5	US-09-833-245-628
21	87.5	24.3	4545	5	US-09-750-972-2
22	86.5	23.9	501	1	PCT-US01-04096A-3346
23	86.5	23.9	4544	1	PCT-US01-04096A-1753
24	85.5	23.6	1139	1	PCT-US01-01302-99
25	85.5	23.6	1139	1	PCT-US01-01302-873
26	85	23.5	43	5	US-09-467-997-5
27	84.5	23.3	288	1	PCT-US01-01302-107

28	84	23.2	2136	1	PCT-US01-04096A-1357	Sequence 1357, Ap
29	83	22.9	527	5	US-09-422-601-1	Sequence 1, Appl
30	82.5	22.8	339	5	US-09-467-997-4	Sequence 106, App
31	82.5	22.8	201	1	PCT-US01-01302-106	Sequence 4, Appl
32	82.5	22.8	379	5	US-09-808-689-4	Sequence 4, Appl
33	82.5	22.8	379	5	US-09-423-844-4	Sequence 4, Appl
34	82.5	22.8	379	5	US-09-403-296A-4	Sequence 4, Appl
35	82.5	22.8	379	5	US-09-380-139A-4	Sequence 3, Appl
36	82	22.7	621	1	PCT-US01-12257-3	Sequence 91, Appl
37	81.5	22.5	148	1	PCT-US01-01302-91	Sequence 101, App
38	81.5	22.5	229	1	PCT-US01-01302-101	Sequence 63, Appl
39	81.5	22.5	3014	1	PCT-US01-10484-63	Sequence 3, Appl
40	81	22.4	41	5	US-09-467-997-3	Sequence 90, Appl
41	81	22.4	213	1	PCT-US01-01302-90	Sequence 1599, Ap
42	81	22.4	213	1	PCT-US01-01321-1559	Sequence 14, Appl
43	79	21.8	534	5	US-09-804-156-14	Sequence 1402, Ap
44	78.5	21.7	323	1	PCT-US01-11988-1402	
45	78.5	21.7	323	5	US-09-833-245-1402	

ALIGNMENTS

```
RESULT 1
US-09-828-366-21
: Sequence 21, Application US/09828366
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Goddard, Audrey
: APPLICANT: Gueney, Austin L.
: APPLICANT: Klein, Robert D.
: APPLICANT: Napier, Mary
: APPLICANT: Wood, William I.
: APPLICANT: Yuan, Jean
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
: FILE OF INVENTION: CELL GROWTH
: FILE REFERENCE: P1694R1C1
: CURRENT APPLICATION NUMBER: US/09/828,366
: CURRENT FILING DATE: 2001-04-05
: Prior filing data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 29
: SEQ ID NO 21
: LENGTH: 723
: TYPE: PRT
: ORGANISM: Homo Sapien
US-09-828-366-21

Query Match          31.2%  Score 113;  DB 5;  Length 723;
Best Local Similarity 52.5%;  Pred. No. 5.5e-06;
Matches 21;  Conservative 3;  Mismatches 14;  Indels 2;  Gaps 1;

OY 24 DICDPNCGNGICGLGLAVGSFSCGCPDGTDPNCSVV 63
DB 445 DDCASSPCANGICRG--VNDSCCTCPRTGNCAPV 482

RESULT 2
US-09-249-697A-7
: Sequence 7, Application US/09249697A
: GENERAL INFORMATION:
: APPLICANT: Ford, John
: APPLICANT: Yeung, George
: TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL
: TITLE OF INVENTION: LIVER SLEEN
: FILE REFERENCE: 24011-727
: CURRENT APPLICATION NUMBER: US/09/249,697A
: CURRENT FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: US 08/966,800
: PRIOR FILING DATE: 1997-11-22
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 7
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(42)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-249-697A-7
```

```
Query Match          30.1%; Score 109; DB 5; Length 42;
Best Local Similarity 36.2%; Pred. No. 1e-06;
Matches 18; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
```

```
QY      24 DICDPNCPENGICLPGLAVGSPFCPCPDGFT 55
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      2 DECSXNCPGNGTCXXDXVGSYXCXCPPGFT 33
```

```
RESULT      3
PCT-US01-12012-3
; Sequence 3, Application PC/TUS0112012
; GENERAL INFORMATION:
; APPLICANT: Genesance Pharmaceuticals
; APPLICANT: Chew, Anne
; APPLICANT: Choi, Julie
; APPLICANT: Koshiy, Beena
; TITLE OF INVENTION: Haplotypes of the F9 Gene
; FILE REFERENCE: MMW-138PCT F9
; CURRENT APPLICATION NUMBER: PCT/US01/12012
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/196,419
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapien
PCT-US01-12012-3
```

```
Query Match          29.1%; Score 105.5; DB 1; Length 461;
Best Local Similarity 36.2%; Pred. No. 2.6e-05;
Matches 21; Conservative 6; Mismatches 20; Indels 11; Gaps 2;
```

```
QY      2 KRSAVWLVLTGSLGVPQSGKDIDPNPCENGICLPGLAVGSPFCPCPDGFT 59
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      82 ERTTEFW-----KQYVDGDCESNFCNLGSSCKDD--INSYECMCFEGEKNC 128
```

```
RESULT      4
US-09-467-997-1
; Sequence 1, Application US/09467997
; GENERAL INFORMATION:
; APPLICANT: Kitajewski, Jan
; APPLICANT: Uytendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: PRT
; ORGANISM: mouse
US-09-467-997-1
```

```
Query Match          28.3%; Score 102.5; DB 5; Length 1964;
Best Local Similarity 44.2%; Pred. No. 0.00023;
```

```
Matches 19; Conservative 2; Mismatches 11; Indels 11; Gaps 2;
```

```
QY      24 DICDPNCPENGIC-----LPGLAVGSPFCPCPDGFTDNCS 60
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      1044 DLSQSPCSNGSGCEITGTGPPG-----FTCHCPKFGPGTCS 1081
```

```
RESULT      5
US-60-248-505-791
; Sequence 791, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
```

```
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: C1000918
```

```
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 791
; LENGTH: 4044
; TYPE: PRT
```

```
; ORGANISM: Human
US-60-248-505-791
```

```
Query Match          27.6%; Score 100; DB 6; Length 4044;
Best Local Similarity 50.0%; Pred. No. 0.00089;
Matches 17; Conservative 4; Mismatches 9; Indels 4; Gaps 1;
```

```
QY      26 CDPNCPENGICLPGLAVGSPFCPCPDGFTDPNC 59
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      1594 CDSPCKNSGFC---SMGSPFCDDCPVGFSGKDC 1623
```

```
RESULT      6
PCT-US01-01302-103
; Sequence 103, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ07PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01302-103
```

```
Query Match          27.1%; Score 98; DB 1; Length 310;
Best Local Similarity 41.7%; Pred. No. 0.00013;
Matches 15; Conservative 6; Mismatches 15; Indels 0; Gaps 0;
```

```
QY      24 DICDPNCPENGICLPGLAVGSPFCPCPDGFTDPNC 59
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB      101 DPCEKNKNGCNGTCVAMAMGKATCRGASGFTGDC 136
```

```
RESULT      7
PCT-US01-01302-89
; Sequence 89, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ07PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
```


Wed May 23 06:29:21 2001

us-09-237-981-14_copy_33_97.rapn

Page 4

;; PRIOR FILING DATE: 2000-04-25
;; NUMBER OF SEQ ID NOS: 2267
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 770
;; LENGTH: 383
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US01-11988-770

Query Match 25.1%; Score 91; DB 1; Length 383;
Best Local Similarity 45.9%; Pred. No. 0.001;
Matches 17; Conservative 3; Mismatches 15; Indels 2; Gaps 1;

OY 24 DICDPNCGGICLPGLAVGSFSCPCPDGFTDPNCS 60
DB 176 DDCLMRPCANGATCLDG--INRFSCLCPEGFAGRFT 210

RESULT 12
US-09-833-245-770
;; Sequence 770, Application US/09833245
;; GENERAL INFORMATION:
;; APPLICANT: Human Genome Sciences, Inc.
;; TITLE OF INVENTION: Albumin Fusion Proteins
;; FILE REFERENCE: PF546PCT
;; CURRENT APPLICATION NUMBER: US/09/833,245
;; PRIOR FILING DATE: 2001-04-12
;; PRIOR APPLICATION NUMBER: 60/229, 358
;; PRIOR FILING DATE: 2000-04-12
;; PRIOR APPLICATION NUMBER: 60/256, 931
;; PRIOR FILING DATE: 2000-12-21
;; PRIOR APPLICATION NUMBER: 60/199, 384
;; PRIOR FILING DATE: 2000-04-25
;; NUMBER OF SEQ ID NOS: 2267
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 770
;; LENGTH: 383
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-833-245-770

Query Match 25.1%; Score 91; DB 5; Length 383;
Best Local Similarity 45.9%; Pred. No. 0.001;
Matches 17; Conservative 3; Mismatches 15; Indels 2; Gaps 1;

OY 24 DICDPNCGGICLPGLAVGSFSCPCPDGFTDPNCS 60
DB 176 DDCLMRPCANGATCLDG--INRFSCLCPEGFAGRFT 210

RESULT 13
PCT-US01-01302-92
;; Sequence 92, Application PC/TUS0101302
;; GENERAL INFORMATION:
;; APPLICANT: Human Genome Sciences, Inc., et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PT207PCT
;; CURRENT APPLICATION NUMBER: PCT/US01/01302
;; PRIOR FILING DATE: 2001-01-17
;; NUMBER OF SEQ ID NOS: 192
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 92
;; LENGTH: 226
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (133)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-92

Query Match 24.9%; Score 90; DB 1; Length 226;
Best Local Similarity 41.0%; Pred. No. 0.0008;
Matches 16; Conservative 3; Mismatches 18; Indels 2; Gaps 1;

OY 26 CDPNCGGICLPGLAVGSFSCPCPDGFTDPNCS 64
DB 28 COSNPLNNVACED--QVGGFMCKCPGFLGTGCKNVD 64

RESULT 14
US-09-802-704-11
;; Sequence 11, Application US/09802704
;; GENERAL INFORMATION:
;; APPLICANT: Godbole, Shubhada
;; APPLICANT: Boyle, Bryan J
;; APPLICANT: Arterburn, Matthew C
;; APPLICANT: Liu, Chenghua
;; APPLICANT: Tang, Y Tom
;; APPLICANT: Dimnac, Radoje T
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AN
;; FILE REFERENCE: HYS-40
;; CURRENT APPLICATION NUMBER: US/09/802,704
;; PRIOR FILING DATE: 2001-03-08
;; PRIOR APPLICATION NUMBER: US 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 22
;; SOFTWARE: Patentin version 3.0
;; SEQ ID NO 11
;; LENGTH: 48
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-802-704-11

Query Match 24.3%; Score 88; DB 5; Length 48;
Best Local Similarity 44.1%; Pred. No. 0.00031;
Matches 15; Conservative 3; Mismatches 14; Indels 2; Gaps 1;

OY 26 CDPNCGGICLPGLAVGSFSCPCPDGFTDPNC 59
DB 16 CSSOPCNGGTCVCG--VNOYRCICPPGRTGNRC 47

RESULT 15
US-09-802-704-7
;; Sequence 7, Application US/09802704
;; GENERAL INFORMATION:
;; APPLICANT: Godbole, Shubhada
;; APPLICANT: Boyle, Bryan J
;; APPLICANT: Arterburn, Matthew C
;; APPLICANT: Liu, Chenghua
;; APPLICANT: Tang, Y Tom
;; APPLICANT: Dimnac, Radoje T
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO FIBULIN-LIKE POLYPEPTIDES AN
;; FILE REFERENCE: HYS-40
;; CURRENT APPLICATION NUMBER: US/09/802,704
;; PRIOR FILING DATE: 2001-03-08
;; PRIOR APPLICATION NUMBER: US 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 22
;; SOFTWARE: Patentin version 3.0
;; SEQ ID NO 7
;; LENGTH: 415
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-802-704-7

Wed May 23 06:29:21 2001

us-09-237-981-14_copy_33_97.rpn

Page 5

Query Match 24.3%; Score 88; DB 5; Length 415;
Best Local Similarity 44.1%; Pred. No. 0.0024;
Matches 15; Conservative 3; Mismatches 14; Indels 2; Gaps 1;
QY 26 CDNRCENGICLRLGLAVGSFSCCECPDGFDPNC 59
DB 116 CSSQPCONGJCVG--VMDTRCICPPGRTGRC 147

Search completed: May 23, 2001, 06:13:55
Job time: 149 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:20 ; Search time 38.64 Seconds
(without alignments)
32.316 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_97
Perfect score: 362
Sequence: 1 MKRSVAVMLVGLSLGVPDQ.....FSCGCPDGFDPNCSSVVEV 65

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/p/odata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/p/odata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/p/odata/2/iaa/5A.COMB.pep:*
4: /cgn2_6/p/odata/2/iaa/5B.COMB.pep:*
5: /cgn2_6/p/odata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/p/odata/2/iaa/Dackfilltest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	362	100.0	513	2 US-08-480-229C-14	Sequence 14, Appl
2	362	100.0	513	2 US-08-480-229C-14	Sequence 14, Appl
3	304	84.0	221	2 US-08-480-229C-29	Sequence 29, Appl
4	304	84.0	221	2 US-08-480-229C-29	Sequence 29, Appl
5	304	84.0	480	2 US-08-480-229C-10	Sequence 10, Appl
6	304	84.0	480	2 US-08-480-229C-10	Sequence 10, Appl
7	222.5	61.5	57	2 US-08-480-229C-23	Sequence 23, Appl
8	222.5	61.5	57	2 US-08-480-229C-23	Sequence 23, Appl
9	127	35.1	25	2 US-08-480-229C-22	Sequence 22, Appl
10	127	35.1	25	2 US-08-480-229C-22	Sequence 22, Appl
11	121	33.4	1404	2 US-08-611-729A-2	Sequence 2, Appl
12	121	33.4	1404	2 US-08-611-729A-2	Sequence 2, Appl
13	117	32.3	2556	1 US-08-185-432-17	Sequence 17, Appl
14	117	32.3	2556	1 US-08-185-432-17	Sequence 17, Appl
15	117	32.3	2556	1 US-08-590A-20	Sequence 20, Appl
16	117	32.3	2556	1 US-08-590A-20	Sequence 20, Appl
17	113	31.2	385	1 US-08-597-545-1	Sequence 1, Appl
18	113	31.2	385	1 US-08-597-545-1	Sequence 1, Appl
19	113	31.2	2523	1 US-08-457-432-18	Sequence 18, Appl
20	112	30.9	383	1 US-08-597-545-2	Sequence 2, Appl
21	112	30.9	383	1 US-08-597-545-2	Sequence 2, Appl
22	111	30.7	157	4 US-08-872-855-6	Sequence 6, Appl
23	111	30.7	157	4 US-08-872-855-6	Sequence 6, Appl
24	109	30.1	713	4 US-08-872-855-5	Sequence 5, Appl
25	109	30.1	720	4 US-08-872-855-4	Sequence 4, Appl
26	109	30.1	1065	2 US-08-400-159-8	Sequence 8, Appl
27	109	30.1	1148	4 US-08-882-046-4	Sequence 4, Appl
			1248	4 US-08-882-046-6	Sequence 6, Appl

28	109	30.1	1257	3 US-08-611-729A-8	Sequence 8, Appl
29	105.5	29.1	461	3 US-08-742-877-2	Sequence 2, Appl
30	105.5	29.1	461	6 US-08-742-877-2	Sequence 2, Appl
31	105	29.0	415	1 US-08-073-531B-1	Sequence 1, Appl
32	105	29.0	415	1 US-08-295-411-2	Sequence 2, Appl
33	105	29.0	415	2 US-08-955-471-2	Sequence 2, Appl
34	105	29.0	415	2 US-08-955-471-2	Sequence 2, Appl
35	105	29.0	415	5 PCT-US92-10242-2	Sequence 1, Appl
36	103	28.5	729	4 US-08-872-855-8	Sequence 8, Appl
37	103	28.5	2471	1 US-08-185-432-16	Sequence 16, Appl
38	103	28.5	2471	1 US-08-083-590A-19	Sequence 19, Appl
39	103	28.5	2471	3 US-08-532-384-19	Sequence 19, Appl
40	103	28.5	2703	1 US-08-185-432-19	Sequence 19, Appl
41	102	28.0	735	3 US-09-191-647-9	Sequence 9, Appl
42	101.5	28.0	721	4 US-08-872-855-7	Sequence 7, Appl
43	101	27.9	52	3 US-08-978-741-15	Sequence 15, Appl
44	101	27.9	448	5 PCT-US92-10068-1	Sequence 1, Appl
45	101	27.9	1193	2 US-08-400-159-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-480-229C-14
; Sequence 14, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-480-229C-14

Query Match 100.0%; Score 362; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 4,Je-32;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGGLAVGSFSCPCPDGFTDNC 60
Db 33 MKRSVAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGGLAVGSFSCPCPDGFTDNC 92
QY 61 SYVEV 65
Db 93 SYVEV 97

RESULT 2

US-08-659-235C-14
; Sequence 14, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-659-235C-14

Query Match 100.0%; Score 362; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 4,1e-32;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGGLAVGSFSCPCPDGFTDNC 60
Db 33 MKRSVAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGGLAVGSFSCPCPDGFTDNC 92
QY 61 SYVEV 65
Db 93 SYVEV 97
RESULT 3
US-08-480-229C-29
; Sequence 29, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:

APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-480-229C-29

Query Match 84.0%; Score 304; DB 2; Length 221;
Best Local Similarity 84.6%; Pred. No. 3,6e-26;
Matches 55; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGGLAVGSFSCPCPDGFTDNC 60
Db 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCPENGICLPGGLAVGSFSCPCPDGFTDNC 60
QY 61 SYVEV 65
Db 61 SYVEV 65

RESULT 4
US-08-659-235C-29
; Sequence 29, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States

RESULT 5
US-08-480-229C-10
Sequence 10, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quertemous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
City: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 516
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462

```

GENERAL INFORMATION:
APPLICANT: Quartermoun, Thomas
APPLICANT: Hogan, Bridg
APPLICANT: Stodermas, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penrite & Edmonds LLP
STREET: 115 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Penrite
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235C-10

```

Query Match 84.0%; Score 304; DB 2; Length 480;

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/480,229C
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-0026-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 Pennle
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 25 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-480-229C-22

Query Match 35.1%; Score 127; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKRSVAVMLVGLSLGVPOFGKDI 25
DB 1 MKRSVAVMLVGLSLGVPOFGKDI 25

RESULT 10
US-08-659-235C-22
Sequence 22, Application US/08659235C
Patent No. 5877281
GENERAL INFORMATION:
APPLICANT: Quartermoun, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown

;; MOLECULE TYPE: peptide
US-08-659-235C-22
Query Match 35.1%; Score 127; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKRSVAVMLVGLSLGVPOFGKDI 25
DB 1 MKRSVAVMLVGLSLGVPOFGKDI 25

RESULT 11
US-08-400-159-2
Sequence 2, Application US/08400159
Patent No. 5869282
GENERAL INFORMATION:
APPLICANT: Ish-Horowitz, David
APPLICANT: Henrique, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Wyatt, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Altvantis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,159
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1404 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-400-159-2

Query Match 33.4%; Score 121; DB 2; Length 1404;
Best Local Similarity 60.0%; Pred. No. 2.2e-05;
Matches 21; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

OY 26 CDRPCEGNGICPLGLAVGSEFCPCPDGFTDPNCS 60
DB 799 CSPPCNGGICIDG--DGFCECMGSGWTGRKCS 831

RESULT 12
US-08-611-729A-2

```
; Sequence 2, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myac, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-611-729A-2

Query Match          33.4%; Score 121; DB 3; Length 1404;
Best Local Similarity 60.0%; Pred. No. 2.2e-05;
Matches 21; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

OY 26 CDPNPGGICLPGLAVGSPFCPCPDGFTDPNCS 60
DB 799 CSPNRCNGICLDG-DGDTCTCMGCGTKRCS 831

RESULT 13
; US-08-185-432-17
; Sequence 17, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
```

```
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2556 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-185-432-17

Query Match          32.3%; Score 117; DB 1; Length 2556;
Best Local Similarity 50.0%; Pred. No. 0.00011;
Matches 20; Conservative 4; Mismatches 14; Indels 2; Gaps 1;

OY 24 DICPNPCNGGICLPGLAVGSPFCPCPDGFTDPNCS 63
DB 660 DECANPCNCHNGICEDG--INGFTCRCPGTHDPTCLSEV 717

RESULT 14
; US-08-083-590A-20
; Sequence 20, Application US/08083590A
; Patent No. 5786158
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,590A
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 20:
```



```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2556 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
;
US-08-083-590A-20

```

```

Query Match 32.3%; Score 117; DB 1; Length 2556;
Best Local Similarity 50.0%; Pred. No. 0.00011;
Matches 20; Conservative 4; Mismatches 14; Indels 2; Gaps 1;

```

```

OY 24 DICDPNCPENGICLPGIAGVSFSCPCPDGFTDPNCSSV 63
DB 680 DEAGNPNCHNGTCEIDG--INGFTCRCPGEGYHPTCLSEV 717

```

RESULT 15

```

US-08-532-384-20
; Sequence 20; Application US/08532384
; Patent No. 6083904
; GENERAL INFORMATION:
; APPLICANT: Attavanis-Teakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,384
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/083,590
; FILING DATE: 25-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2556 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
;
US-08-532-384-20

```

```

Query Match 32.3%; Score 117; DB 3; Length 2556;
Best Local Similarity 50.0%; Pred. No. 0.00011;
Matches 20; Conservative 4; Mismatches 14; Indels 2; Gaps 1;

```

```

OY 24 DICDPNCPENGICLPGIAGVSFSCPCPDGFTDPNCSSV 63
DB 680 DEAGNPNCHNGTCEIDG--INGFTCRCPGEGYHPTCLSEV 717

```

Search completed: May 23, 2001, 06:13:21
Job time: 115 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:12:44 ; Search time 62.53 Seconds
(without alignments)
446.210 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Perfect score: 2264
Sequence: 1 GPCTPNPNCHNGSTCEISRAY.....MSWYGRITVLASELLCTEE 406

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1095	48.4	409	2	T11743
2	1081	47.7	427	2	JC4915
3	1075	47.5	463	1	A36479
4	1072	47.3	401	2	S65138
5	1072	47.3	427	2	S74211
6	682.5	30.1	2183	2	T42764
7	670.5	29.6	2224	1	KFHUS
8	670	29.6	2319	2	A47004
9	657.5	29.0	2211	1	KFBOS
10	649	28.7	2351	1	E2HU
11	643	28.4	218	2	A47285
12	639	28.2	2133	2	T42763
13	464.5	20.5	216	2	A44258
14	437	19.3	927	1	JQ0948
15	368.5	16.3	3133	2	S52093
16	262.5	11.6	845	2	JC5256
17	248	11.0	2352	2	T30201
18	234	10.3	2318	2	S45306
19	230	10.2	2524	2	A35844
20	229.5	10.1	2531	2	S18188
21	228.5	10.1	2531	2	A46019
22	228	10.1	2437	2	S42612
23	228	10.1	2531	2	T31070
24	226	10.0	1064	2	A40136
25	225	9.9	719	2	S51739
26	224.5	9.9	2703	1	A24420
27	223.5	9.9	728	2	S150719
28	223	9.8	2471	2	A49128
29	221	9.8	1203	2	A49175

30	220.5	9.7	861	2	A48825	Notch homolog Motc
31	220	9.7	1331	2	T42218	slit-1 protein hom
32	220	9.7	2555	2	A40043	notch protein homo
33	214.5	9.5	2321	2	S78549	notch3 protein - h
34	213	9.4	570	2	A48836	fibropellin C prec
35	209.5	9.3	293	2	B26637	neurogenic repelit
36	209.5	9.3	2139	2	A35672	crumbs protein - f
37	208	9.2	1408	2	S16148	gene serrate prote
38	208	9.2	1523	2	T13953	MEGF5 protein - ra
39	207.5	9.2	1220	2	A56136	jagged protein pre
40	203.5	9.0	1283	2	T13799	neurexin IV - frui
41	200.5	8.9	832	2	A31246	neurogenic protein
42	200.5	8.9	833	2	S19087	gene Delta protein
43	200.5	8.9	880	2	S06670	neurogenic repelit
44	200	8.8	1025	2	T42626	secreted leucine-r
45	198	8.7	1687	2	T30176	EGF repeat transme

ALIGNMENTS

RESULT 1
T11743
pp47 protein - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 16-Jul-1999 #sequence, revision 16-Jul-1999 #text, change 21-Jul-2000
C:Accession: T11743
R:Ensslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toe
Biol. Reprod. 58, 1057-1064, 1998
A:Title: Molecular cloning and characterization of P47, a novel boar sperm-associated
A:Reference number: Z17325; M0JD:98206817
A:Accession: T11743
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-409 <ENS>
A:Cross-references: EMBL:Y11683; NID:92652927; PIDN:CAA72379.1; PID:92652928
A:Experimental source: testis
C:Function:
A:Description: may be involved in membrane remodeling and/or function as a zona pelli
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:6-40/Domain: EGF homology <EGF>

Query Match	48.48	Score 1095	DB 2	Length 409
Best Local Similarity	50.18	Pred. No. 5.6e-76		
Matches 203	Conservative 60	Mismatches 98	Indels 44	Gaps 5
QY	1	GPCTPNPNCHNGSTCE-ISEAYRGDTFLGYVCKCRGFNGIHCOHNIKEVEPECKNGIC	59	
DB	46	GPCFPNPNCHNDACEVIDAHRGDVFEYICKCPHGVTGIH	86	
QY	60	TDLVAVYSECEGFEFMGRNCOYKCSGLIEGGIISNOQITASSSTRHALFGLOKWPYYA	119	
DB	87	-----CEIICNAPLQKMETGALIDFQISASSMHLQFMELQWABELA	127	
QY	120	RLMKGLINAMTAENDRWKRWIOINLORRMVYGVITGAKRISPEYIKFYIAVSND	179	
DB	128	RLRRAGIVAAWTAWSYDR-NPMIOVINLIRRMVYGVITGASRASASPMYKFFVAASD	186	
QY	180	GKTMAYKVKKTNE--DWVFRGNIDNNTPYANSFTPIPAQVRLYPQVCRHRTLMEL	237	
DB	187	GRFQF--IQGAESGDKIFPMGNIDNSGLKVNLEFVLEVVRLVPIICRGCTLRFEL	244	
QY	238	LGSELSCGSEPLGMSGHIOYOTASSIFRTLMNMTFWERKARLQKQVAAWMSGH	297	
DB	245	LGSELSCGSEPLGMDNTIPKQITASSFRTGSLAFSWFFYARLQNKQFNAWMAOS	304	
QY	298	NDOSQWLQVDLVPKRVYGVITGQAKDFGVQVSGYKLAYSNDGEHWTVYQDEKQKDK	357	
DB	305	NSASEWLQIDLSQRRVGVITGQAKDFGHIOYVAAYKVAYSDDSVSTEVYRDQALGK	364	
QY	358	VFGNFDNDTHRKNVIDPPIYARIRILPMSWGRITVLASELLGC	402	

DB 365 IFPGMLDNNSHKKNMFEPLTRPVRLPVAHNRITLRELLLOC 409

RESULT 2

aga protein precursor - rat
N/Alternate names: O-acetyl-Gd3 ganglioside
C/Species: Rattus norvegicus (Norway rat)
C/Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 20-Jun-2000
C/Accession: J04915
R/Ogura, K.; Kura, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sana, Y.
Biochem. Biophys. Res. Commun. 225, 932-938, 1996
A/Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.
A/Reference number: J04915; MUID:96374422
A/Accession: J04915
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-427 <OCU>
A/Cross-references: DDBJ:DB4068; NID:g1620006; PIDN:BAAL220.1; PID:g1620007
A/Experimental source: C57 cell
C/Comment: This protein is required for the O-acetylation of disialoganglioside sialic acid.
A/Genes: ags
C/Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homology
F/1-21/Domain: signal sequence #status predicted <Sig>
F/28-60/Domain: EGF homology <EG1>
F/68-107/Domain: EGF homology <EG2>
F/110-267/Domain: discoidin I amino-terminal homology <DN1>
F/271-427/Domain: discoidin I amino-terminal homology <DN2>

Query Match 47.7%; Score 1081; DB 2; Length 427;
Best local similarity 50.4%; Pred. No. 6.8e-75;
Matches 208; Conservative 59; Mismatches 120; Indels 26; Gaps 7;

QY 3 CTPNCHNGGCEISEAYRBDTFIGYCKPRGFGNGHCOINME-----CEVEPKNGK 57
DB 28 CQSSICLNGGICLMDQ---DNDI--YCLCPREFGLVC---NETKGCSPNPFHDA 77
QY 58 ICT-----DLVANSCECPGEFMRNCOYKCSGPLIEGIIISNOQITASSSTHRALE 109
DB 78 KCLVETDRTORGDIFTEYICOCPVGYSGICELGCEGKLGGAIDASQISASVYMGPM 137
QY 110 GLOKRYPYARANKKGLINAMTAENDRMKRWIQLNLRKRVGVITOGAKRIGSEPT 169
DB 138 GLOHMGPELARLYRTGIVNMTASSYDS--KPMIQLVDFLRKRVGCVTQGSARAGAEYL 196
QY 170 KEKTIAYNSNDGKTAMKVKGTNEDMVFNGNIDNMTFYANSFTPIKAQVVRXPOVCR 229
DB 197 KTFKVAISUDGRREFEIQDESIGDKEFMGNODNNSLKIMNFNPTLEAQITRLPVSCHR 256
QY 230 HCTLMELLGCELSGCEPLGMSGHIQDYQITASSIFRTLNMDFTWEPKRALDKQK 289
DB 257 GCTLRFELGCELGCESEPLGLKNNTIPDSQITASSYKTNLRAFWYDHLRDLNCK 316
QY 290 VNAVMSGNHDQSOVLQVLDLVPTKVTGITTQCAKDFGVQVGSYKLAYSNDGEHWTVY 349
DB 317 INAWTAQNSAKENWLDLTGQKVGITITQCARDFGHQVYVAYSKVAHSDGVQTVY - 375
QY 350 DEKQKDKVFGQGNFNDTRKKNVYDPIYARHRIILPMSWGRITLASELLGC 402
DB 376 -EEQSTKSVFGQNLNNSHKKNIFEKPFMARVRYVLPVSWHNRITLRELLGC 427

RESULT 3

milk fat globule membrane protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: A36479
R/Stubbbs, J.D.; Iekutis, C.; Singer, K.L.; Bul, A.; Yuzuki, D.; Stinivasan, U.; Parry, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990
A/Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the ex

A/Reference number: A36479; MUID:91046008

A/Accession: A36479
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-463 <STU>
A/Cross-references: GB:M38337; NID:g199142; PIDN:AAA9534.1; PID:g199143
C/Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homology
F/28-60/Domain: EGF homology <EG1>
F/68-107/Domain: EGF homology <EG2>
F/147-303/Domain: discoidin I amino-terminal homology <DN1>
F/307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match 47.5%; Score 1075; DB 1; Length 463;
Best local similarity 50.1%; Pred. No. 2.1e-74;
Matches 203; Conservative 69; Mismatches 123; Indels 10; Gaps 6;

QY 1 GCPNCHNGGCEISEAYRBDTFIGYCKPRGFGNGHCOINMECEVEPKNGK 59
DB 66 GPCSPNCHNGGCEISEAYRBDTFIGYCKPRGFGNGHCOINMECEVEPKNGK 122
QY 60 TDLYANSCCEPGEF--MGRNCOYKCSGPLIEGIIISNOQITASSSTHRALEGLDKMTPY 117
DB 123 TPAVPAVPAVPTPDPDLNNSLASRSTQLMEGALADSOISASTVYMGGLDRMPE 182
QY 118 YARLNKGLINAMTAENDRMKRWIQLNLRKRVGVITOGAKRIGSEPTIKFYKIAS 177
DB 183 LARLYRTGIVNMTASSYDS--KPMIQLVDFLRKRVGCVTQGSARAGAEYLKTFKVA 241
QY 178 NDGKTAMKVKGTNEDMVFNGNIDNMTFYANSFTPIKAQVVRXPOVCRHCTLRMEL 237
DB 242 LDGRKFEIIDE--SGDKEFLGNDNNSLKIMNFNPTLEAQITRLPVSCHRGCTLRREL 300
QY 238 LGCELSGCEPLGMSGHIQDYQITASSIFRTLNMDFTWEPKRALDKQGVNMTSGH 297
DB 301 LQCEHGLCEPLGKKNWITPDSQITASSYKTNLRAFWYDHLRDLNCKINMTAQ 360
QY 298 NDQSOVLQVLDLVPTKVTGITTQCAKDFGVQVGSYKLAYSNDGEHWTVYDEKQKDK 357
DB 361 NSAKEMQLVLDLTGQKVGITITQCARDFGHQVYVAYSKVAHSDGVQTVY--EEQSSK 418
QY 358 VFQGNFNDTRKKNVYDPIYARHRIILPMSWGRITLASELLGC 402
DB 419 VFQGNLNNSHKKNIFEKPFMARVRYVLPVSWHNRITLRELLGC 463

RESULT 4

glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
N/Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFC-E
C/Species: Bos primigenius taurus (cattle)
C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-Aug-1998
C/Accession: S65138; MUID:96125736
R/Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochem. Biophys. Acta 1245, 385-391, 1995
A/Title: Molecular cloning of glycoprotein antigen MGP57/53 recognized by monoclonal
A/Reference number: S65138; MUID:96125736
A/Accession: S65138
A/Status: preliminary
A/Molecule type: not compared with conceptual translation
A/Residues: 1-401 <AKK>
R/Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A/Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-
ii-like sequences.
A/Reference number: A48394; MUID:93250576
A/Accession: G48394
A/Status: preliminary
A/Molecule type: protein
A/Residues: 207-220 <MAT>
A/Experimental source: milk
A/Note: sequence extracted from NCBI backbone (NCBIP.131457)

Db 1919 KPMIOVMOKEVVYVVTGIOTOGANKHYLKSCFTTEFOVAYSSDOTWMOIFRGKSGKSVMYET 1978
 QY 199 GNIDNTPYANSFPTPIKAQVYLYPOVCRHCTRLRMELLCGELSCGSEPLGMSGHIOD 258
 Db 1979 GNSDGTSTIKENRLDPPVAVYIRIHPTKSYNRPTRLRLDLOGCEVNGCSTPLGLEDEGHIO 2038
 QY 259 YQIASSIFITLMDMFTNEPRKARLDKQKVNAMTSGHNDQSOQLVOYLVPKVTGII 318
 Db 2039 KQIYASSFKKSMWGDY--WEPSLARLNAGRYNANAAKANNKQMLDVLTKIKRYAIY 2096
 QY 319 TQAGKDEGHVQVGSYKLYASNDGEMHTVYODEKORAKVEGFDNDTHKKNVIDPPYI 378
 Db 2097 TQCGKSLSEEMVYKSYSIGQSDGVAMKRYRQKSSMVDKIEFGNSNTRGHKKNFNPPII 2156
 QY 379 ARHRIPLMPSMGRITLASELLGC 402
 Db 2157 SRIRIRIPKTMNOSIALRLLEFGC 2180

RESULT 7
 REF05
 coagulation factor V precursor (validated) - human
 N:Alternate names: coagulation labile factor; proaccelerin
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1989 #sequence revision 02-Jun-1995 #text_change 08-Dec-2000
 R:Clpep: L.D.; Moore, K.D.; Kane, W.H.
 Biochemistry 31, 3777-3785, 1992
 A:Title: Structure of the gene for human coagulation factor V.
 A:Reference number: A42344; MUID:92232668
 A:Accession: A56172
 A:Molecule type: DNA
 A:Residues: 1-2224 <CR>
 A:Cross-references: GB:M1785
 A:Accession: A42344
 A:Molecule type: DNA
 A:Residues: 48-56;79-89;120-130;191-201;229-249;313-323;368-378;428-437;461-471;533-542;
 2070;2111-2120;2172-2181 <CR>
 R:Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kaufm
 Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
 A:Title: Complete cDNA and derived amino acid sequence of human factor V.
 A:Reference number: A28028; MUID:87260886
 A:Accession: A28028
 A:Molecule type: mRNA
 A:Residues: 1-857 'R', 859-864, 'R', 866-924, 'E', 926-1763, 'I', 1765-2212, 'T', 2214-2224 <EN>
 A:Cross-references: GB:M16967
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete
 R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davle, E.W.
 Biochemistry 26, 6508-6514, 1987
 A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of hum
 A:Reference number: A27498; MUID:88107560
 A:Accession: A27498
 A:Molecule type: mRNA
 A:Residues: 1-1284, 'I', 1286-1600 <KAN>
 A:Cross-references: GB:M17785
 A:Note: parts of this sequence were determined by protein sequencing
 R:Kane, W.H.; Davle, E.W.
 Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986
 A:Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homolog
 A:Reference number: A25897; MUID:86313665
 A:Accession: A25897
 A:Molecule type: mRNA
 A:Residues: 1188-1215, 1315-2224 <KA2>
 A:Cross-references: GB:M14335
 A:Note: parts of this sequence were determined by protein sequencing
 R:Keller, F.G.; Ortel, T.L.; Quinn-Alten, M.A.; Kane, W.H.
 Biochemistry 34, 4118-4124, 1995
 A:Title: Thrombin-catalyzed activation of recombinant human factor V.
 A:Reference number: A56139; MUID:95210278
 A:Contents: annotation: thrombin cleavage sites
 C:Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
 C:Genetics:
 A:Gene: GDB:f5

A:Cross-references: GDB:119896; OMIM:227400
 A:Map position: 1q23-1q23
 A:Insertions: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1.
 A:Deletion:
 A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa p
 A:Pathway: blood coagulation
 C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-2224/Product: coagulation factor V #status predicted <MAT>
 F:29-737/Product: coagulation factor Va heavy chain #status experimental <VAH>
 F:29-345/Domain: A1 <DB1>
 F:339-329/Domain: ferroxidase repeat homology <FO1>
 F:346-691/Domain: A2 <DA2>
 F:351-684/Domain: ferroxidase repeat homology <FO2>
 F:692-1573/Domain: B <DOB>
 F:1183-1461/Region: 9-residue repeats (Q-X-1/N-L-S-P-D-L-S)
 F:1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>
 F:1574-1905/Domain: A3 <DA3>
 F:1581-1905/Domain: ferroxidase repeat homology <FO3>
 F:1667-1765/Region: phospholipid binding #status predicted
 F:1906-2064/Domain: C1 <DC1>
 F:1906-2061/Domain: discoidin I amino-terminal homology <DNI>
 F:2065-2224/Domain: C2 <DC2>
 F:2065-2221/Domain: discoidin I amino-terminal homology <DN2>
 F:31,35,239,297,460,468,534,741,752,760,776,782,821,938,977,1074,1083,1103,1106,1479,
 F:157-193,248-329,500-526,603-684,1723-1751,1907-2061,2066-2221/disulfide bonds: #sta
 F:334-335/Cleavage site: Arg-Asn (protein C) #status predicted
 F:363,693,1546/Binding site: sulfate (Tyr) (covalent) #status predicted
 F:376-777/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
 F:382,1338/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:534-535/Cleavage site: Arg-Gly (protein C) #status predicted
 F:737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
 F:1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experime
 F:1573-1574/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 29.6%; Score 670.5; DB 1; Length 2224;
 Best Local Similarity 41.8%; Pred. No. 8.2e-43;
 Matches 141; Conservative 54; Mismatches 127; Indels 15; Gaps 4;

QY 69 ECGEEMGRNCQYKSCGPIGIGISNOITASTHRAFGLOKWPYARLRKGLN 128
 Db 1897 QTFLIMDKRCR-----PGLSTGIISDQTKASEF-----LGIWBPRLRINNGSYD 1946

QY 129 AMTA--AENDRWKRWIQINLQKRWTVGTQAKRIGSEYIKFYKTIAYSDGKTAM 185
 Db 1947 AWSVEKLAIEFASKPWIOYDMQEVIIIGIOTOGANKHYLKSCFTTEFYAYSSNOIMQI 2006

QY 186 YKKGTEEMVERGNIDNNTPYANSFPTPIKAQVYLYPOVCRHCTRLRMELLCGELSC 245
 Db 2007 FKNSRNNMYNNGNSDASTIKENQFPPYAKYIRISPRATYNNRPTLRLELOGCEVNGC 2066

QY 246 SEPLGKSGHIDYQITVASSIFRILNMDFTNEPRKARLDKQKVNAMTSGHNDQSOQL 305
 Db 2067 STEPLGMEKNIENKQITASSFKSMWGDY--WEPSRARLNAGRYNANAAKANNKQML 2124

QY 306 VDLVPTKVTGIIOTGARDGCHVQVGSYKLYASNDGEMHTVYODEKORAKVEGFDND 365
 Db 2125 IDLTKIKTKITAIITIGCKSLSEMYVKNYTIHYSEGVENKPRRLSSMWDKIFEBSNTWT 2184

QY 366 DTHRKNVIDPPYAKYIRILMPSMGRITLASELLGC 402
 Db 2185 KGHVKNFNPPIISFRIRIVIPKTMNOSIALRLLEFGC 2221

RESULT 8
 A47004
 coagulation factor VIII precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
 A:Accession: A47004
 R:Elider, B.; Laksch, D.; Glitscher, J.

Hum. Mol. Genet. 1, 199-200, 1992
A:Title: Sequence of the exon-containing regions of the human factor VIII gene.
A:Reference number: 154318; MUID:93265012
A:Accession: 154318
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1921, 'S', 1923-2351 <RES>
A:Cross-references: GB:M88648; NID:q182381; PIDN:AAA52420.1; PID:q182383
R:Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschler, J.; Keyt, B.; Seeburg
Nature 312, 330-337, 1984
A:Title: Expression of active human factor VIII from recombinant DNA clones.
A:Reference number: A00525; MUID:85061548
A:Accession: A00525
A:Molecule type: mRNA
A:Residues: 1-2351 <MOO>
A:Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
R:Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Soltzman, L.A.; Buecker, J.L.; Pittman, D.D.;
S, D.N.; Hewick, R.M.
Nature 312, 342-347, 1984
A:Title: Molecular cloning of a cDNA encoding human antihemophilic factor.
A:Reference number: 158059; MUID:85061550
A:Accession: 158059
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RE2>
A:Cross-references: GB:K01740; NID:q182802; PIDN:AAA52484.1; PID:q182803
R:Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; Kuo,
B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.; No
DNA 4, 333-349, 1985
A:Title: Characterization of the polypeptide composition of human factor VIII:C and the
A:Reference number: A23584; MUID:86081164
A:Accession: A23584
A:Molecule type: mRNA
A:Residues: 1-2351 <TRU>
A:Cross-references: GB:M14113; NID:q182817; PIDN:AAA52485.1; PID:q182818
R:Edton, D.; Rodriguez, H.; Vehar, G.A.
Biochemistry 25, 505-512, 1986
A:Title: Proteolytic processing of human factor VIII. Correlation of specific cleavages
17.
A:Reference number: A26174; MUID:86159740
A:Accession: A26174
A:Molecule type: protein
A:Residues: 20-36;336-399, 'X', 401-402;1668-1678;1709-1722, 'D', 1723-1725;1741-1755 <EAT>
R:Pittman, D.D.; Wang, J.H.; Kautman, R.J.
Biochemistry 31, 3315-3325, 1992
A:Title: Identification and functional importance of tyrosine sulfate residues within re
A:Reference number: A42348; MUID:92207952
A:Accession: A42348
A:Molecule type: protein
A:Residues: 20-36;336-371;392-408;582-594;1668-1669, 'X', 1671;1672-1692;1693-1708;1709-17
A:Experimental source: recombinant material from Chinese hamster ovary cells
A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
R:Fay, P.J.; Smudzyn, T.M.
J. Biol. Chem. 264, 14005-14010, 1989
A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
A:Reference number: A43986; MUID:89340500
A:Accession: A43986
A:Molecule type: protein
A:Residues: 'X', 517-523;1853-1860, 'X', 1862-1864, 'X', 1866 <RAY>
R:Leyte, A.; van Schijndel, H.B.; Niehs, C.; Hutter, W.B.; Verbeet, M.P.; Mertens, K.;
J. Biol. Chem. 266, 740-746, 1991
A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
A:Reference number: A56109; MUID:91093266
A:Contents: annotation; sulfation
R:Gitschler, J.; Wood, W.I.; Goralka, T.M.; Wilson, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
Nature 312, 326-330, 1984
A:Title: Characterization of the human factor VIII gene.
A:Reference number: A56196; MUID:85061547
A:Contents: annotation; Introns
R:Komulien, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
Protein Sci. 4, 740-746, 1995
A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains of
A:Reference number: A56216; MUID:95338127

A:Contents: annotation; disulfide bonds
A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
R:Kjalke, M.; Heding, A.; Taiho, G.; Persson, E.; Thomsen, J.; Ezban, M.
Eur. J. Biochem. 234, 773-779, 1995
A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
A:Reference number: S63527; MUID:96163459
A:Accession: S63527
A:Molecule type: protein
A:Residues: 733-752;753-759 <KJA>
R:Lind, P.; Larsson, K.; Spira, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandb
Eur. J. Biochem. 232, 19-27, 1995
A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construct
A:Reference number: S66445; MUID:96048024
A:Accession: S66445
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1668-1685 <LIN>
C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure
C:Genetics:
A:Gene: GDB:F8C
A:Cross-references: GDB:119124; OMIM:306700
A:Map position: Xq28-Xq28
A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3;
C:Function:
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa
A:Pathway: blood coagulation
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxi
C:Keywords: acute phase, blood coagulation; duplication; glycoprotein; hemophilia A;
F:10-19/Domain: signal sequence [status predicted <SIG>
F:20-2351/Product: coagulation factor VIII [status experimental <MAT>
F:20-740/Product: coagulation factor VIIa heavy chain [status experimental <ACH>
F:20-356/Domain: A1 <DA1>
F:23-348/Domain: ferroxidase repeat homology <FO1>
F:392-759/Domain: A2 <DA2>
F:402-730/Domain: ferroxidase repeat homology <FO2>
F:1668-2251/Product: coagulation factor VIIa light chain [status experimental <ACL>
F:1709-2038/Domain: A3 <DA3>
F:1716-2038/Domain: ferroxidase repeat homology <FO3>
F:2039-2191/Domain: C1 <DC1>
F:2039-2188/Domain: discoidin I amino-terminal homology <DN1>
F:2192-2351/Domain: C2 <DC2>
F:2192-2345/Domain: discoidin I amino-terminal homology <DN2>
F:610,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,140
F:1172-1398,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #sta
F:335-336/Cleavage site: Arg-Met (coagulation factor Xa, protein C) [status predicted
F:365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) [status experiment
F:391-393/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) [status experiment
F:414,426/Binding site: sulfate (Tyr) (covalent) [status predicted
F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) [status experiment
F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) [status experimental
F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) [status experime
F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) [status experimental
F:2193-2345/Disulfide bonds: #status predicted

Query Match 28.7%; Score 649; DB 1; Length 2351;
Best Local Similarity 44.2%; Pred. No. 3.8e-41;
Matches 144; Conservative 46; Mismatches 120; Indels 16; Gaps 5;

QY 82 KCSGLGIEBGIIISNOQTASTSTHAFGLQKWPYRARKLKKGLINAWTAENDRRKRW 141
|| ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 2039 KQPLPLGASGHIRDFQITASQY-----SQMAPKLRLHYSSINAWSTKEP--FSW 2089
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 142 IOINQRMRYTGVITQCAKRGSPYIKFKIAYNSNGKTMWYKKGCTMEDVDFRNI 201
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 2090 IKVDLAPMIHIGTKQCARQRFSSLYTSQITIMSLDGKMWQYRNGNSTGLTWVDFGNV 2149
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 202 DNNPYANSFTPIKAQYVRYPQVCRHRCYLRMELGCELSGCSPEPLGKMSGHIQDYOI 261
|| : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 2150 DSSGKIHNIENPPIARIRLHPHYISRSRLRMELMOCDLNCSMPLGMSKAIISDAOI 2209
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 262 TASSIFRILNMDMFWPEPRARLDKQCKVNAWMTSGHNDOSQMLQVLDLLVPTKYTGIIITQ 320

Db 2210 TASSYF---TMMFATWSPSKARHLQGRSNAMRPQVNNPKEMLOVDFOKTKVGTQTQ 2265
 QY 321 GAKDGHVQVFGSYKLAASNDGEHMTVYODEKQKDKVFGQGFNDTHRKVNIDPPIYAR 380
 Db 2266 GKSLTSMYVEEFLISSQDGHQWTLF--FONGKVKVFGQGDSTFPPVNSLDPPLLTR 2323
 QY 381 HIRILPMSWYGRITLASELLGCTEE 406
 Db 2324 YLRHPQSWVHQIALRMEVLGCEAOD 2349

RESULT 11

A47285
 milk fat globule protein - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 28-May-1999
 C:Accession: A47285
 R:Latocca, D.; Peterson, J.A.; Ureia, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
 Cancer Res. 51, 4994-4998, 1991
 A:Title: A-M-r 46,000 human milk fat globule protein that is highly expressed in human b
 A:Reference number: A47285; MUID:91373351
 A:Accession: A47285
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-218 <LAR>
 A:Cross-references: GB:S56151; NID:9235396; PIDN:AAH19771.1; PID:9235397
 C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
 F:1-56/Domain: discoidin I amino-terminal homology (fragment) <DN1>
 F:60-218/Domain: discoidin I amino-terminal homology <DN2>

Query Match 28.4%; Score 643; DB 2; Length 218;
 Best Local Similarity 56.8%; Pred. No. 6,6e-42;
 Matches 117; Conservative 30; Mismatches 59; Indels 0; Gaps 0;
 QY 197 FGNIDNNTPVANSFTPIKAQVRLYQVCRHCTLRMELIGELSCSEPLGKSGHI 256
 Db 13 FVGNNKNAHVNLFEFEVQVRLYPTSCHTACTLRFEELGELGNCANPLGIKNSI 72
 QY 257 ODYOTASSIFRTLMDETPPRKARLDKQKVNAMTSGHNDOSQWLQVLDLVPRTYTG 316
 Db 73 PKQITASSSYTQGLHLSMNPSTARDKQGFENAWAGSYGNDQWLQVLDGSSKEVTG 132
 QY 317 IITQAKDFGHVQVFGSYKLAASNDGEHMTVYODEKQKDKVFGQGFNDTHRKVNIDP 376
 Db 133 IITQAKRNGSQVYASVYKVAASNDASAWTEQDPRQTSKIFPGNMDSHKKNLFEETP 192
 QY 377 IYARHRIILPMSWYGRITLASELLG 402
 Db 193 ILARYRILIPVAMHNRILRLLELLGC 218

RESULT 12

142763
 coagulation factor VIII precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
 C:Accession: T42763
 R:Lollar, P.
 Submitted to the EMBL Data Library, August 1996
 A:Reference number: Z2269
 A:Accession: T42763
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2133 <LO>
 A:Cross-references: EMBL:U49517; NID:91511633; PID:91511634; PIDN:AAH06705.1
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pig
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
 F:23-349/Domain: ferroxidase repeat homology <FOX1>
 F:402-730/Domain: ferroxidase repeat homology <FOX2>

F:148-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 28.2%; Score 639; DB 2; Length 213;
 Best Local Similarity 40.0%; Pred. No. 1.9e-40;
 Matches 144; Conservative 52; Mismatches 136; Indels 28; Gaps 6;

QY 47 ECEPEPKNGICIDLVANSCPEGFEMGRNCOYKSSGLIEGIIISMOQTASTHR 106
 Db 1800 ECLIGELHQAGMSTTFLV-YSKS-----CQALPGMASGRIRDFQITAGQY 1844
 QY 107 ALFGLQKWPYPAFLNKKGLINMTAENDRWKRWIOINLQRMRYTGVITQAKRIGSP 166
 Db 1845 -----GQWAPKRLANLHSGSINAMSTKDP---HSWIKVDLLAMIIHGIMTQARQKFS 1896
 QY 167 EYIFKVIANSNDGKTWAMTKVGTNEDWYFRGNDINNTPVANSFTPIKAQVRLYPOV 226
 Db 1897 LYISQFLIMYSLDGRNMQSYRGNSTGTLMWFCGNVDSGIRKNIIPNPVARYIRLPHN 1956
 QY 227 CRRHCTLRMELIGELSCSEPLGKSGHIQDYOTASSIFRTLMDETPPRKARLDK 286
 Db 1957 YSISTRLMELMGCDLNSCMPLOMKAISDSOITASS--HLSNIFATWSPQARLH 2013
 QY 287 QGKVNAMTSGHNDOSQWLQVLDLVPRTYTGIIITQAKDFGHVQVFGSYKLAASNDGEHMT 346
 Db 2014 QGRTNMRPRVSSAEWLOVDLQKTVKVTGITTQGVKSLSSYKVEFLVSSQDGRMT 2073
 QY 347 VYODEKQKDKVQVGNNDNTHRKVNIDPPIYARHRIILPMSWYGRITLASELLGCTEE 406
 Db 2074 LFIQDGH--TKVFGQNDSSSTPPVNALDPLFLTRYLRHPTSWAHOIALRLMEVLGCEAOD 2131

RESULT 13

A44258
 factor VIII-associated gene B hypothetical protein - human
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
 C:Accession: A44258
 R:Levinson, B.; Kenwick, S.; Gamel, P.; Fisher, K.; Gitscher, J.
 Genomics 14, 585-589, 1992
 A:Title: Evidence for a third transcript from the human factor VIII gene.
 A:Reference number: A44258; MUID:93052386
 A:Accession: A44258
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-216 <LEV>
 A:Cross-references: GB:M90707; NID:9182316; PIDN:AAA58466.1; PID:9182317
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid
 F:1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>
 F:57-210/Domain: discoidin I amino-terminal homology <DN2>

Query Match 20.5%; Score 464.5; DB 2; Length 216;
 Best Local Similarity 46.7%; Pred. No. 2.5e-28;
 Matches 99; Conservative 31; Mismatches 75; Indels 7; Gaps 3;

QY 196 VERNIDNNTPVANSFTPIKAQVRLYQVCRHCTLRMELIGELSCSEPLGKSGH 255
 Db 9 VFGNVDSGKIHFIENPILARIIRLPHYSIRSLRLELMGCDLNSCMPLGMSRA 68
 QY 256 IODYOTASSIFRTLMDETFEPRKARLDKQKVNAMTSGHNDOSQWLQVLDLVPRTY 314
 Db 69 ISDAQITASSYF---TMMFATWSPSKARHLQGRSNAMRPQVNNPKEMLOVDFOKTKMV 124
 QY 315 TGIITQAKDFGHVQVFGSYKLAASNDGEHMTVYODEKQKDKVFGQGFNDTHRKVNID 374
 Db 125 TGVYTOQVKSILTSMYKEFLISSQDGHQWTLF--FONGKVKVFGQNDSTFPPVNSLD 182
 QY 375 PTYARHRIILPMSWYGRITLASELLGCTEE 406
 Db 183 PPLLTRILRHPQSWVHQIALRMEVLGCEAOD 214

MEGM_MOUSE STANDARD; PRT; 463 AA.

AC P21956; P97800; 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-BGF FACTOR 8) (MFG-E8) (MFGM)

GN MFG-E8.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-35.

RC TISSUE=Mammary gland;

RC MEDLINE=91046008; PubMed=2122462;

RA Stubbs J.D., Iekulis C., Singer K.L., Bul A., Yuzuki D., Srivivasan U., Parry G.;

RT "cDNA cloning of a mouse mammary epithelial cell surface protein reveals the existence of epidermal growth factor-like domains linked to factor VIII-like sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).

RN [2]

RP SEQUENCE OF 23-463 FROM N.A.

RC TISSUE=Testis;

RA Ensslin M.A.;

RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-BINDING PROTEIN.

CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN

CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND SPERMATOZOAN.

CC -1- DEVELOPMENTAL STAGE: MRNA EXPRESSION IS DETECTABLE IN MAMMARY TISSUE FROM NONPREGNANT ANIMALS & MAXIMAL IN THE LACTATING GLAND.

CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M38337; AAA39534.1; -

DR EMBL; Y11684; CAAT2380.1; -

DR PIR; A36479; A36479.

DR HSSP; P00740; 11XA.

DR MGD; MGI:102768; Mfeg8.

DR InterPro; IPR000421; -

DR InterPro; IPR000561; -

DR InterPro; IPR001438; -

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00754; F5_F8_Type_C; 2.

DR PRINTS; PRO0010; EGFBLD.

DR PROSITE; PS00022; EGF_1; 2.

DR PROSITE; PS01186; EGF_2; 2.

DR PROSITE; PS01285; FAS5C_1; 2.

DR PROSITE; PS01286; FAS5C_2; 2.

DR Signal; Glycoprotein; Repeat; EGF-like domain; MILK.

FT SIGNAL 1 22

FT CHAIN 23 463

FT GN 24

FT DOMAIN 24 61

FT DOMAIN 64 108

FT DOMAIN 148 303

FT DOMAIN 308 463

FT SITE 87 89

FT DISULFID 28 39

FT DISULFID 33 49

FT DISULFID 51 60

FT DISULFID 68 79

FT BY SIMILARITY.

FT DISULFID 73 96

FT DISULFID 98 107

FT DISULFID 148 303

FT DISULFID 290 294

FT DISULFID 308 463

FT CARBOHYD 61 61

FT CARBOHYD 266 266

FT CARBOHYD 316 316

FT CARBOHYD 426 426

FT CARBOHYD 30 30

FT CONFLICT 35 35

FT CONFLICT 110 147

FT CONFLICT 168 168

FT CONFLICT 196 196

FT CONFLICT 309 309

FT CONFLICT 395 395

SO SEQUENCE 463 AA; 51465 MM; D78B6C6EFBA724D CRC64;

Query Match 47.5%; Score 1075; DB 1; Length 463;

Best Local Similarity 50.1%; Pred. No. 5 3e-78;

Matches 203; Conservative 69; Mismatches 123; Indels 10; Gaps 6;

QY 1 GPCPNPCNGTCEIS-EAYRQDTPLGYCKCPGFGNHCQHNINECEVEPCNGGIC 59

DB 66 GPCSPNCPYNDKVLITDORGIDFLEYICCPVSGIHCTETNYIND--GEYMF 122

QY 60 PLVAVNYSCCEPGF--MGRNCQKCGPLGICGIIISMOITASTTRALFGLQKWPY 117

DB 123 TTAVENTAVPPATPPLDLSNNLSRCSIOGMGAGAIASQISAVYVMGMLQRMPE 182

QY 118 YARLKKGLINAMTAENDNRKRIQINLQKKRVGTVTQCAKIGSPETIKTKIAYS 177

DB 183 IARLRTGIVANAMASVND-KPIQVNLRLKRVAGVMTQASAGAEYIKTFKVAYS 241

QY 178 NDGKTWAKYKKGNTEDMVEFGNIDNNTPYANSFTPIKAQVRLYPQVCRHRTLMEL 237

DB 242 LGRFEFTIOE-SGSKPEFLGNLNNLSLKVNMENFTLEQVIRLYPSCHRGCTLREL 300

QY 238 LCCESLGGSEPLGKSGHIDYQTASTSIFRLNMMFTWEPKRLRLQKQKVAMVSGH 297

DB 301 LCCELHGLGELPLGKNNITPDSQMSASSKYTNLAFNMYHLGRLDQKINAMVQS 360

QY 298 NDQSMLOVDLLPLTKVTGIIITQAKDFHVPFVSGYKLAYSNDSEHMTYODEKQRDK 357

DB 361 NSAKEMLOVDLGTOROVYGIITQGRDFGHIOYVESKYVAHSDGVQWTVY--EEOGSSK 418

QY 358 VFQGNFNDTRKKNVIDPPIYARHRLPMSWYGRITLASELLGC 402

DB 419 VFQGLDNNSHKKNIFKFPFARVYKLPVSWNRITLLELLGC 463

RESULT 4

MEGM_BOVIN STANDARD; PRT; 427 AA.

AC 095114; Q27959; P79344; 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-BGF FACTOR 8) (MFG-E8)

DE (MG957/53) (PAS-6/PAS-7) GLYCOPROTEIN (MFGM) (SPERM SURFACE PROTEIN SP47) (BP47) (COMPONENTS 15/16).

GN MFG-E8.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.

OX NCBI_Taxid=9913;

RN [1]

RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.

RC STRAIN=HOLSTEIN; TISSUE=Mammary gland;

RC MEDLINE=97008954; PubMed=8856064;

RA Hvarregaard J., Andersen M.H., Berglund L., Rasmussen J.T.,
 RA Petersen T.E.;
 RA "Characterization of glycoprotein PAS-6/7 from membranes of bovine
 RL milk fat globules";
 RL Eur. J. Biochem. 240:628-636(1996).
 RN (2)
 RN SEQUENCE OF 18-427 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=96125736; PubMed=8541316;
 RA Aoki N., Kishi M., Taniguchi Y., Adachi T., Nakamura R.,
 RA Matsuda T.;
 RT "Molecular cloning of glycoprotein antigens MGp57/53 recognized by
 RT monoclonal antibodies raised against bovine milk fat globule
 RL membrane";
 RL Biochim. Biophys. Acta 1245:385-391(1995).
 RN (3)
 RN SEQUENCE OF 19-427 FROM N.A.
 RC TISSUE=Testis;
 RA Ensslin M.A.;
 RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RP [4]
 RP SEQUENCE OF 140-146; 174-187; 233-246 AND 422-427.
 RC TISSUE=Milk;
 RX MEDLINE=93250576; PubMed=8485470;
 RA Mather I.H., Banghart L.R., Lane W.S.;
 RT "The major fat globule membrane proteins, bovine components 15/16 and
 RT guinea-pig GP 35, are homologous to MGF-E8, a murine glycoprotein
 RT containing epidermal growth factor-like and factor V/VIII-like
 RT sequences.";
 RL Biochem. Mol. Biol. Int. 29:545-554(1993).
 CC -1- FUNCTION: PROBABLY ASSOCIATES WITH PHOSPHOLIPIDS ON THE SURFACE OF
 CC MAMMARY EPITHELIAL CELLS AND MILK FAT GLOBULES. ZONA PELLUCIDA-
 CC BINDING PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM
 CC LACKS 53 AMINO ACIDS WITHIN THE FS/8 TYPE C 1 DOMAIN.
 CC -1- TISSUE SPECIFICITY: MILK AND SPERMATOZOAN.
 CC -1- PTM: THE 2 O-LINKED GLYCANS CONSIST OF GAL, GLCNAC AND FUC. WITH
 CC PROBABLY FUC AS REDUCING TERMINAL SUGAR.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 FS/8 TYPE C DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X91895; CAA62997.1; -
 DR EMBL; S80643; AAB35894.2; -
 DR EMBL; Y11719; CAA72406.1; -
 DR HSSP; P00740; IIXA.
 DR InterPro: IPR000421; -
 DR InterPro: IPR000561; -
 DR Pfam; PF00754; F5_F8_type_C_2.
 DR Pfam; PF00008; EGF_2.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS0186; EGF_2; 2.
 DR PROSITE; PS01285; FA58C_1; 2.
 DR PROSITE; PS01286; FA58C_2; 2.
 DR Signal; Glycoprotein; Milk; Repeat; EGF-like domain;
 KW Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 427
 FT DOMAIN 20 59
 FT DOMAIN 62 106
 FT DOMAIN 109 265
 FT DOMAIN 270 427
 FT SITE 85 87
 FT DISULFID 24 35
 FT DISULFID 29 47
 FT BY SIMILARITY.

FT DISULFID 49 58 BY SIMILARITY.
 FT DISULFID 66 77 BY SIMILARITY.
 FT DISULFID 71 94 BY SIMILARITY.
 FT DISULFID 96 105 BY SIMILARITY.
 FT DISULFID 109 265
 FT DISULFID 252 256
 FT DISULFID 270 427
 FT CARBOHYD 27 27
 FT CARBOHYD 34 34
 FT CARBOHYD 59 59
 FT CARBOHYD 227 227
 FT VARSPLIC 169 221
 FT CONFLICT 19 19
 FT CONFLICT 28 28
 SQ SEQUENCE 427 AA; 47411 MW; 4CBBE3A1DC4EB24 CRC64;
 O-LINKED (FUC. . .) (IN PAS-6).
 O-LINKED (FUC. . .) (IN PAS-7).
 N-LINKED (GLCNAC. . .) (HYBRID) (IN PAS-6
 AND PAS-7).
 N-LINKED (GLCNAC. . .) (HIGH MANNOSE) (IN
 PAS-6).
 MISSING (IN SHORT ISOFORM).
 A -> F (IN REF. 1).
 L -> Q (IN REF. 1).
 Query Match 47.3%; Score 1072; DB 1; Length 427;
 Best Local Similarity 47.9%; Pred. No. 8,3e-78;
 Matches 193; Conservative 68; Mismatches 102; Indels 40; Gaps 3;
 QY 1 GPCPNPCNCHGTCGCEIS-EAYRGDFEIGYCKCPRGFGNGIHCHONINECEVEPCKNKGIC 59
 DB 64 GPCPNPCNCHDACEQYTDSDSHRGDVFIOYICKCPGIVGHCERT----- 108
 QY 60 TDLVANYSCCEPGEPEMGKRCQYKCSGPGIGIEGIIISNOOITPASTHRALEFGLOKWTPIYA 119
 DB 109 -----CTSPLEMOGTGAIAOISOISASSMHIGEMGLORAPBELA 145
 QY 120 RLNKKGLINAMTAENDRKRMIOINLORKMRYGVITOGAKRIGSEYIKFYKIAYNSD 179
 DB 146 RLHGTGVNMTSGNDK-NPMIYVNLKRMVYGVITOGASRAGSEYIKFEKAYSTSD 204
 QY 180 GKTAAKRYVGTEDNAPFGNDNTNTPANSPTPIKAQVYRLXPQVCRHCTRLMELIG 239
 DB 205 GRQFQFIOVAGRSDDITFGVANNVNSGLKINLEDPLEQYRLVLPITICHGCTRELIG 264
 QY 240 CELSGCSEPLMGKSGHIDYOITASSIFRTLMDMFTWEPKARLKOGRVNAWTSGHND 299
 DB 265 CELNGCTEPLGLKDNITIPNKQITASYYKWTGLSAFSEFPYARLDNOGFNMTAQWTS 324
 QY 300 QSQMLQVLDLVPKRYGIIITOGAKDFGHQVFGSKLAYNSDEHMTYVODEKQKRDYVF 359
 DB 325 ASENLDLDSQKRYGIIITOGARDFGHQVAYARAYGDDGYTWEYKDPQASSTKIF 384
 QY 360 QGNFNDTHRKNIYDPIYARHRIILPMSWTGRITLASELIGC 402
 DB 385 PQNNDNNSHKKNIEFTFPQARFVRIQPVAMHNRITLREVELIGC 427
 RESULT 5
 MFGM_HUMAN STANDARD; PRT: 387 AA.
 ID MFGM_HUMAN 008431;
 AC 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (HMF8)
 DE (BRASR EPITHELIAL ANTIGEN BA46) (MFGM) [CONTAINS: MEDIN].
 GN MFG8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCIT_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast, and Breast carcinoma;
 RX MEDLINE=96213908; PubMed=8639284;
 RA Couto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Petersen J.A.;
 RT "Cloning and sequence analysis of human breast epithelial antigen
 RT BA46 reveals an RGD cell adhesion sequence presented on an epidermal


```

RL Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987).
RN [3]
RP SEQUENCE OF 1-1600 FROM N.A.
RX MEDLINE-88107560; PubMed-2827731;
RA Kane W.H., Ichihose A., Hagen F.S., Davie E.W.;
RT "Cloning of cDNAs coding for the heavy chain region and connecting
RT region of human factor V, a blood coagulation factor with four types
RT of internal repeats."
RL Biochemistry 26:6508-6514(1987).
RN [4]
RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.
RX MEDLINE-86313665; PubMed-3092220;
RA Kane W.H., Davie E.W.;
RT "Cloning of a cDNA coding for human factor V, a blood coagulation
RT factor homologous to factor VIII and ceruloplasmin."
RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).
RN [5]
RP PARTIAL SEQUENCE FROM N.A.
RC TISSUE-fibroblast;
RX MEDLINE-93203619; PubMed-8454869;
RA Shen N.L.L., Fan S.-T., Pyati J., Graff R., Lapolla R.J.,
RT Edgington T.S.;
RT "The serine protease cofactor factor V is synthesized by
RT lymphocytes."
RL J. Immunol. 150:2992-3001(1993).
RN [6]
RP VARIANT APCR GLN-534.
RX MEDLINE-94217810; PubMed-8164741;
RA Bertina R.M., Koelman B.P.C., Koster T., Rosendaal F.R.,
RA Dieren R.J., de Ronde H., van der Velden P.A., Reijnders P.H.;
RT "Mutation in blood coagulation factor V associated with resistance to
RT activated protein C."
RL Nature 369:64-67(1994).
CC -1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES
CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN
CC -1- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT
CC CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS
CC IS CALCIUM-DEPENDENT.
CC -1- DOMAIN: DOMAIN B CONTAINS 35 X 9 AA TANDEN REPEATS. AND 2 X 17 AA
CC REPEATS.
CC -1- PTM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE
CC COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-
CC TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).
CC -1- DISEASE: OMRENN PARAHEMOPHILIA, AN HEMORRHAGIC DIAPYCNOSIS, IS DUE
CC TO A DEFICIENCY OF FACTOR V. OTHER DEFECTS IN F5 RESULTS IN A
CC FORM OF THROMBOPHILIA KNOWN AS APC RESISTANCE (APCR). THE APCR
CC MUTATION IS FOUND IN ABOUT 5% OF THE POPULATION WHICH SUGGEST THAT
CC A SLIGHT THROMBOTIC TENDENCY MAY CONFER SOME ADVANTAGE IN FETAL
CC IMPLANTATION.
CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
CC 2 PLASTOCYANIN-LIKE REPEATS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L32779; AAB59401.1; JOINED.
DR EMBL; L32755; AAB59401.1; JOINED.
DR EMBL; L32756; AAB59401.1; JOINED.
DR EMBL; L32757; AAB59401.1; JOINED.
DR EMBL; L32758; AAB59401.1; JOINED.
DR EMBL; L32759; AAB59401.1; JOINED.
DR EMBL; L32760; AAB59401.1; JOINED.
DR EMBL; L32761; AAB59401.1; JOINED.
DR EMBL; L32762; AAB59401.1; JOINED.
DR EMBL; L32763; AAB59401.1; JOINED.
DR EMBL; L32764; AAB59401.1; JOINED.
DR EMBL; L32765; AAB59401.1; JOINED.
DR EMBL; L32766; AAB59401.1; JOINED.
DR EMBL; L32767; AAB59401.1; JOINED.
DR EMBL; L32768; AAB59401.1; JOINED.
DR EMBL; L32769; AAB59401.1; JOINED.
DR EMBL; L32770; AAB59401.1; JOINED.
DR EMBL; L32771; AAB59401.1; JOINED.
DR EMBL; L32772; AAB59401.1; JOINED.
DR EMBL; L32773; AAB59401.1; JOINED.
DR EMBL; L32774; AAB59401.1; JOINED.
DR EMBL; L32775; AAB59401.1; JOINED.
DR EMBL; L32776; AAB59401.1; JOINED.
DR EMBL; L32777; AAB59401.1; JOINED.
DR EMBL; L32778; AAB59401.1; JOINED.
DR EMBL; M16967; AAB59401.1; JOINED.
DR EMBL; M14335; AAB59532.1; JOINED.
DR PIR; A25897; A25897.
DR PIR; A28028; A28028.
DR HSSP; P00450; 1KCM.
DR MIM; 134400; -.
DR MIM; 188055; -.
DR MIM; 227310; -.
DR MIM; 227400; -.
DR InterPro; IPR000421; -.
DR InterPro; IPR001117; -.
DR Pfam; PF00754; Cu-oxidase; 3.
DR Pfam; PF00754; F5-F8_type.C; 2.
DR PROSITE; PS00079; MULTICOPEP_OXIDASE1; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR Blood coagulation; Plasma; Glycoprotein; Calcium; Signal; zymogen;
KW Repeat; Polymorphism; Disease mutation; Thrombophilia.
FT FT CHAIN 1 28
FT FT CHAIN 29 2224
FT FT CHAIN 30 737
FT FT PEPTIDE 738 1573
FT FT CHAIN 1574 2224
FT FT DOMAIN 30 329
FT FT DOMAIN 30 193
FT FT DOMAIN 203 329
FT FT DOMAIN 348 684
FT FT DOMAIN 348 526
FT FT DOMAIN 536 684
FT FT DOMAIN 692 1573
FT FT DOMAIN 895 928
FT FT REPEAT 895 911
FT FT REPEAT 912 928
FT FT SIMILAR 1135 1148
FT FT DOMAIN 1185 1463
FT FT REPEAT 1185 1193
FT FT REPEAT 1194 1202
FT FT REPEAT 1203 1211
FT FT REPEAT 1212 1220
FT FT REPEAT 1221 1229
FT FT REPEAT 1230 1238
FT FT REPEAT 1239 1247
FT FT REPEAT 1248 1256
FT FT REPEAT 1257 1265
FT FT REPEAT 1266 1274
FT FT REPEAT 1275 1283
FT FT REPEAT 1284 1292
FT FT REPEAT 1293 1301
FT FT REPEAT 1302 1310
FT FT REPEAT 1311 1319
FT FT REPEAT 1320 1328
FT FT REPEAT 1329 1337
FT FT REPEAT 1338 1346
FT FT REPEAT 1347 1355
FT FT REPEAT 1356 1364
FT FT REPEAT 1365 1373
FT FT REPEAT 1374 1382
FT FT REPEAT 1383 1391
FT FT REPEAT 1391 1391

```


FT REPEAT 1392 1400 24.
 FT REPEAT 1401 1409 25.
 FT REPEAT 1410 1418 26.
 FT REPEAT 1419 1427 27.
 FT REPEAT 1428 1436 28.
 FT REPEAT 1437 1445 29.
 FT REPEAT 1446 1454 30.
 FT REPEAT 1455 1463 31.
 FT REPEAT 1464 1472 32.
 FT REPEAT 1473 1481 33.
 FT REPEAT 1482 1490 34.
 FT REPEAT 1493 1501 35.
 FT DOMAIN 1578 1578 5. F5/8 TYPE A 3.
 FT DOMAIN 1578 1578 5. PLASTOCYANIN-LIKE 5.
 FT DOMAIN 1578 1578 6. PLASTOCYANIN-LIKE 6.
 FT DOMAIN 1578 1578 6. F5/8 TYPE C 1.
 FT DOMAIN 1578 1578 6. F5/8 TYPE C 2.
 FT SITE 1046 1047 738 CLEAVAGE (BY THROMBIN).
 FT SITE 1046 1047 738 CLEAVAGE (BY THROMBIN).
 FT SITE 1573 1574 193 CLEAVAGE (BY THROMBIN).
 FT DISULFID 167 193 PROBABLE.
 FT DISULFID 500 526 PROBABLE.
 FT DISULFID 1125 1751 BY SIMILARITY.
 FT DISULFID 1407 2061 BY SIMILARITY.
 FT DISULFID 2066 2221 BY SIMILARITY.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match

Best Local Similarity 42.1%; Score 675.5; DB 1; Length 2224;
 Matches 142; Conservative 54; Mismatches 126; Indels 15; Gaps 4;

QY 69 ECPGEPGRNCQYKSGPLIEGIIISNOQTASSTHRLGLOKWPYARLANKGLIN 128
 DB 1897 QTPFLIMDDCM---PMGLSTGIISDSQIKASEF-----LQWEPRLARLNGSGYN 1946
 QY 129 AMTA--AENDRMKRWIOLNLRKMRVGVITOGAKRIGSEPTKFKIAYSDGKTAM 185
 DB 1947 AMSVEKLAERASPKWIOVMOKEVITIGTQGAHKHLKSCYTTEFYVASSNOINMO 2006
 QY 186 YKAKGTNEDEVRGNINNTPTANSTPPRIKAOYRLYPVCRHCTLRNELLGCELSGC 245
 DB 2007 FKGNSTRVMYKENGSDASTIKENQDPPYARIRISPTFRANRPLRLLELGCEVNC 2066
 QY 246 SEPLGKSGHIODOYITASSIFRTLMMDFTPEPRKARLDGCGVNAWTSCHNDOSOMIO 305
 DB 2067 STPLGKSGHIODOYITASSIFRTLMMDFTPEPRKARLDGCGVNAWTSCHNDOSOMIO 305
 QY 306 VDLVPTVGTITOGAKDFGHVQVGSYKLAYSNDGEHWTVYDERKORRDKVPGCNEDN 365
 DB 2125 IDLAKIKITAITOGCKSLSSSEMYKSYTHYSQGVEMPRYRKLSSMDKIFEGNTMT 2184
 QY 366 DTRKSNVIDPPYARIRILPWSMVGITLASELLGCG 402
 DB 2185 KGHVKNFEPPLISRTIRVPIKTNQSTIRLRLGCG 2221

RESULT 7

ID PA8.MOUSE STANDARD; PRT: 2319 AA.

AC 006194;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT).
 GN CF8 OR F8C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=93300511; PubMed=8314577;
 RA Elder B., Lakich D., Gitschler J.;
 RL "Sequence of the murine factor VIII cDNA.";
 CC Genomics 16:374-379(1993).
 CC -1- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 CC AS A COPROFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
 CC ACTIVATED FORM, FACTOR IXA.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
 CC 2 PLASTOCYANIN-LIKE REPEATS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 DR EMBL: L05573; AAA37385.1; .
 DR PIR: A47004; A47004.
 DR HSSP: P00451; ICFG.
 DR MGD: MGI:88383; F8.
 DR InterPro: IPR001117; .
 DR InterPro: IPR000421; .
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
 KW Signal; Glycoprotein; Sulfatation.
 FT CHAIN 1
 FT SIGNAL 19
 FT DOMAIN 20 2319
 FT DOMAIN 20 349
 FT DOMAIN 20 399
 FT DOMAIN 207 349
 FT DOMAIN 399 573
 FT DOMAIN 583 730
 FT DOMAIN 760 1640
 FT DOMAIN 1683 2008
 FT DOMAIN 1683 1845
 FT DOMAIN 1853 2008
 FT DOMAIN 2008 2156
 FT DOMAIN 2161 2313
 FT SITE 391 392
 FT SITE 759 760
 FT SITE 1678 1679
 FT SITE 1324 1325
 FT SITE 1640 1641
 FT MOD_RES 367 367
 FT MOD_RES 737 737
 FT MOD_RES 738 738
 FT MOD_RES 742 742
 FT MOD_RES 1669 1669
 FT MOD_RES 1687 1687
 FT MOD_RES 173 199
 FT DISULFID 547 573
 FT DISULFID 1819 1845
 FT DISULFID 2008 2156
 FT DISULFID 2161 2313
 FT CARBOHYD 61 61
 FT CARBOHYD 233 233
 FT CARBOHYD 259 259

 F5/8 TYPE A 1.
 PLASTOCYANIN-LIKE 1.
 PLASTOCYANIN-LIKE 2.
 F5/8 TYPE A 2.
 PLASTOCYANIN-LIKE 3.
 PLASTOCYANIN-LIKE 4.
 B.
 F5/8 TYPE A 3.
 PLASTOCYANIN-LIKE 5.
 PLASTOCYANIN-LIKE 6.
 F5/8 TYPE C 1.
 F5/8 TYPE C 2.
 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 SULFATATION (BY SIMILARITY).
 SULFATATION (BY SIMILARITY).
 SULFATATION (BY SIMILARITY).
 SULFATATION (BY SIMILARITY).
 SULFATATION (REQUIRED FOR VWF BINDING)
 (BY SIMILARITY).
 SULFATATION (BY SIMILARITY).
 PROBABLE.
 PROBABLE.
 BY SIMILARITY.
 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match	29.68;	Score 670;	DB 1;	Length 2319;
Best Local Similarity	46.38;	Pred. No. 4e-45;		
Matches 151;	Conservative 40;	Mismatches 119;	Indels 16;	Gaps 5.

RA Guinto R.R., Esmon C.T., Mann K.G., Macgillivray R.T. ;
 "The complete cDNA sequence of bovine coagulation factor V. ";
 RL J. Biol. Chem. 267:2971-2978(1992).
 CC -1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES
 CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.
 CC -1- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT
 CC CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS
 CC IS CALCIUM-DEPENDENT.
 CC -1- DOMAIN: DOMAIN B CONTAINS 29.5 X 9 AA TANDEM REPEATS, AND 2 X 17
 CC AA REPEATS.
 CC -1- PTM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE
 CC FORM. FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-
 CC TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS)).
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
 CC 2 PLASTOCYANIN-LIKE REPEATS.
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M81440; AAA30512.1; -.
 CC EMBL: M81441; AAA30513.1; -.
 CC HSSP: P00450; 1KCM.
 CC InterPro: IPR001117; -.
 CC InterPro: IPR001117; -.
 CC pfam: PF00754; Cu-oxidase; 3.
 CC pfam: PF00754; F5_F8_type_C; 2.
 CC PROSITE: PS00079; TWINTOPPR_OXIDASE1; 2.
 CC PROSITE: PS01286; FAS8C_2; 2.
 CC Blood coagulation; Plasma; Glycoprotein; Calcium; Signal; Zymogen;
 CC Repeat.
 CC KM
 CC FT SIGNAL. 28
 CC FT CHAIN 29 2211
 CC FT CHAIN 29 741
 CC FT PEPTIDE 742 1564
 CC FT
 CC FT CHAIN 1565 2211
 CC FT DOMAIN 30 327
 CC FT DOMAIN 30 193
 CC FT DOMAIN 203 327
 CC FT DOMAIN 348 686
 CC FT DOMAIN 348 525
 CC FT DOMAIN 535 686
 CC FT DOMAIN 535 1564
 CC FT SIMILAR 899 915
 CC FT DOMAIN 1124 1151
 CC FT DOMAIN 1124 1137
 CC FT REPEAT 1138 1151
 CC FT REPEAT 1188 1453
 CC FT DOMAIN
 CC FT
 CC REPEAT 1188 1196
 CC REPEAT 1197 1205
 CC REPEAT 1206 1214
 CC REPEAT 1215 1223
 CC REPEAT 1224 1232
 CC REPEAT 1233 1241
 CC REPEAT 1242 1250
 CC REPEAT 1251 1259
 CC REPEAT 1260 1268
 CC REPEAT 1269 1277
 CC REPEAT 1278 1286
 CC REPEAT 1287 1295
 CC REPEAT 1296 1304
 CC REPEAT 1305 1313
 CC REPEAT 1314 1322
 CC REPEAT 1323 1331
 CC REPEAT 1332 1340
 CC REPEAT 1341 1349
 CC REPEAT 1350 1358
 CC REPEAT 1359 1367
 CC REPEAT 1368 1376
 CC REPEAT 1377 1385
 CC REPEAT 1386 1394
 CC REPEAT 1395 1403
 CC REPEAT 1404 1412
 CC REPEAT 1413 1421
 CC REPEAT 1422 1430
 CC REPEAT 1431 1439
 CC REPEAT 1440 1448
 CC REPEAT 1449 1457
 CC REPEAT 1458 1466
 CC REPEAT 1467 1475
 CC REPEAT 1476 1484
 CC REPEAT 1485 1493
 CC REPEAT 1494 1502
 CC REPEAT 1503 1511
 CC REPEAT 1512 1520
 CC REPEAT 1521 1529
 CC REPEAT 1530 1538
 CC REPEAT 1539 1547
 CC REPEAT 1548 1556
 CC REPEAT 1557 1565
 CC REPEAT 1566 1574
 CC REPEAT 1575 1583
 CC REPEAT 1584 1592
 CC REPEAT 1593 1601
 CC REPEAT 1602 1610
 CC REPEAT 1611 1619
 CC REPEAT 1620 1628
 CC REPEAT 1629 1637
 CC REPEAT 1638 1646
 CC REPEAT 1647 1655
 CC REPEAT 1656 1664
 CC REPEAT 1665 1673
 CC REPEAT 1674 1682
 CC REPEAT 1683 1691
 CC REPEAT 1692 1700
 CC REPEAT 1701 1709
 CC REPEAT 1710 1718
 CC REPEAT 1719 1727
 CC REPEAT 1728 1736
 CC REPEAT 1737 1745
 CC REPEAT 1746 1754
 CC REPEAT 1755 1763
 CC REPEAT 1764 1772
 CC REPEAT 1773 1781
 CC REPEAT 1782 1790
 CC REPEAT 1791 1799
 CC REPEAT 1800 1808
 CC REPEAT 1809 1817
 CC REPEAT 1818 1826
 CC REPEAT 1827 1835
 CC REPEAT 1836 1844
 CC REPEAT 1845 1853
 CC REPEAT 1854 1862
 CC REPEAT 1863 1871
 CC REPEAT 1872 1880
 CC REPEAT 1881 1889
 CC REPEAT 1890 1898
 CC REPEAT 1899 1907
 CC REPEAT 1908 1916
 CC REPEAT 1917 1925
 CC REPEAT 1926 1934
 CC REPEAT 1935 1943
 CC REPEAT 1944 1952
 CC REPEAT 1953 1961
 CC REPEAT 1962 1970
 CC REPEAT 1971 1979
 CC REPEAT 1980 1988
 CC REPEAT 1989 1997
 CC REPEAT 1998 2006
 CC REPEAT 2007 2015
 CC REPEAT 2016 2024
 CC REPEAT 2025 2033
 CC REPEAT 2034 2042
 CC REPEAT 2043 2051
 CC REPEAT 2052 2060
 CC REPEAT 2061 2069
 CC REPEAT 2070 2078
 CC REPEAT 2079 2087
 CC REPEAT 2088 2096
 CC REPEAT 2097 2105
 CC REPEAT 2106 2113
 CC REPEAT 2114 2122
 CC REPEAT 2123 2131
 CC REPEAT 2132 2140
 CC REPEAT 2141 2149
 CC REPEAT 2150 2158
 CC REPEAT 2159 2167
 CC REPEAT 2168 2176
 CC REPEAT 2177 2185
 CC REPEAT 2186 2194
 CC REPEAT 2195 2203
 CC REPEAT 2204 2212
 CC REPEAT 2213 2221
 CC REPEAT 2222 2230
 CC REPEAT 2231 2239
 CC REPEAT 2240 2248
 CC REPEAT 2249 2257
 CC REPEAT 2258 2266
 CC REPEAT 2267 2275
 CC REPEAT 2276 2284
 CC REPEAT 2285 2293
 CC REPEAT 2294 2302
 CC REPEAT 2303 2311
 CC REPEAT 2312 2320
 CC REPEAT 2321 2329
 CC REPEAT 2330 2338
 CC REPEAT 2339 2347
 CC REPEAT 2348 2356
 CC REPEAT 2357 2365
 CC REPEAT 2366 2374
 CC REPEAT 2375 2383
 CC REPEAT 2384 2392
 CC REPEAT 2393 2401
 CC REPEAT 2402 2410
 CC REPEAT 2411 2419
 CC REPEAT 2420 2428
 CC REPEAT 2429 2437
 CC REPEAT 2438 2446
 CC REPEAT 2447 2455
 CC REPEAT 2456 2464
 CC REPEAT 2465 2473
 CC REPEAT 2474 2482
 CC REPEAT 2483 2491
 CC REPEAT 2492 2500
 CC REPEAT 2501 2509
 CC REPEAT 2510 2518
 CC REPEAT 2519 2527
 CC REPEAT 2528 2536
 CC REPEAT 2537 2545
 CC REPEAT 2546 2554
 CC REPEAT 2555 2563
 CC REPEAT 2564 2572
 CC REPEAT 2573 2581
 CC REPEAT 2582 2590
 CC REPEAT 2591 2599
 CC REPEAT 2600 2608
 CC REPEAT 2609 2617
 CC REPEAT 2618 2626
 CC REPEAT 2627 2635
 CC REPEAT 2636 2644
 CC REPEAT 2645 2653
 CC REPEAT 2654 2662
 CC REPEAT 2663 2671
 CC REPEAT 2672 2680
 CC REPEAT 2681 2689
 CC REPEAT 2690 2698
 CC REPEAT 2699 2707
 CC REPEAT 2708 2716
 CC REPEAT 2717 2725
 CC REPEAT 2726 2734
 CC REPEAT 2735 2743
 CC REPEAT 2744 2752
 CC REPEAT 2753 2761
 CC RE

"Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor.";
 RT J. Biol. Chem. 266:740-746(1991).
 RL [17]
 RN SULFATION
 RP MEDLINE-92207952; PubMed-1554716;
 RX Pitman D.D., Wang J.H., Kaufman R.J.;
 RA "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII.";
 RL Biochemistry 31:3315-3325(1992).
 RN [18]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX MEDLINE-95200924; PubMed-7893714;
 RA Gilbert G.E., Baleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy.";
 RL Biochemistry 34:3022-3031(1995).
 RN [19]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE-91221499; PubMed-1902642;
 RA Gitschler J.;
 RT "The molecular basis of hemophilia A.";
 RL Ann. N.Y. Acad. Sci. 614:83-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE-89088506; PubMed-2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A.";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE-95245332; PubMed-7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A.";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT GLN-2326.
 RX MEDLINE-86235344; PubMed-2012775;
 RA Gitschler J., Wood W.I., Simmen W.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RL mild hemophiliac.";
 RL Science 232:1415-1416(1986).
 RN [13]
 RP VARIANT PRO-2135.
 RX MEDLINE-88096539; PubMed-3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschler J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences.";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT GLN-2228.
 RX MEDLINE-88191889; PubMed-2833855;
 RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides.";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT GLY-291.
 RX MEDLINE-88220354; PubMed-2835904;
 RA Youssoufian H., Wong C., Aronis S., Platakoukis H., Kazazian H.H. Jr.,
 RT Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from GLU->GLY substitution
 RT in exon 7 of the factor VIII gene.";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT CYS-1708.
 RX MEDLINE-89274393; PubMed-2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A.";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RP VARIANT CYS-391.
 RX MEDLINE-90001543; PubMed-2506948;
 RA Shima M., Ware J., Yoshioaka A., Fukui H., Fulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RP VARIANT LEU-189.
 RX MEDLINE-90057680; PubMed-2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A.";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RP VARIANT LEU-2326.
 RX MEDLINE-89197216; PubMed-2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene.";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RP VARIANT HIS-391.
 RX MEDLINE-89264602; PubMed-2498862;
 RA Aral M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372->histidine).";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT CYS-1708.
 RX MEDLINE-90105723; PubMed-2104766;
 RA Aral M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RT Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material positive hemophilia A.";
 RL Blood 75:384-389(1990).
 RN [22]
 RP VARIANT GLN-2228 AND LEU-2326.
 RX MEDLINE-90123183; PubMed-2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RT Mancuso G., Morfini M., de Biasi R., Baudo F., Cardonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent.";
 RL Blood 75:662-670(1990).
 RN [23]
 RP VARIANT CYS-391.
 RX MEDLINE-90329422; PubMed-1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Atani A., Tuddenham E.G.;
 RT "CRM+ hemophilia A due to a missense mutation (372->Cys) at the
 RT internal heavy chain thrombin cleavage site.";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RP VARIANTS PHE-1699 AND CYS-1708.
 RX MEDLINE-90152691; PubMed-2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RT Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA.";
 RL Genomics 6:65-71(1990).
 RN [25]
 RP VARIANTS CYS-1728 AND ASP-1941.
 RX MEDLINE-90169988; PubMed-2106480;
 RA Tirasman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RT Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene.";
 RL Query Match 28.7%; Score 649; DB 1; Length 2351;
 RL Best Local Similarity 44.2%; Pred. No. 1,9e-43;


```

DB 1800 ECLGELHQAQSMSTLTV-YSKE-----CQAPLGMASGRIRDOQIATSGY- 1844
QY 107 ALFLOLQWYPIYARLKKGLINMTAENDRKRMTQIOLQKKMYVTITOGAKRIGSP 166
DB 1845 -----GGWAPPLALAHSGSINAMSTKDP---HSWKYDLAPMFIHIMTOGAROKRSS 1896
QY 167 EYIKFYKIAVNDGKTAMKVKGTNEDVYFKNIDNTPYANSTPPIKQAYRLXPQV 226
DB 1897 LYISQFIIMYSLDGRNMQSYRGNSGTLMVFGVNDASIKININIPPIVARIIRLHPTH 1956
QY 227 CRRHCTLMELGCELSGCEPILGAKSGHIQDYQITASSIFTLNMDFTPEPKARLDK 286
DB 1957 YSIRSTLMELMCDLNSCPLGQNKALISDSQITASS---HLSNIFATWSPSOARHL 2013
QY 287 QGKVNATSGHNDOSQWLYVDLVPKVTGIIITOGAKDFGHVQFSGYKLAYSNDGEHMT 346
DB 2014 QGRTNAMRPVYSAEEMLYQVDLQKTYKVTGITTQYKSLSSMYKPELVSSQDGRMT 2073
QY 347 VYODEKOKKRVQGNFNDNTHKKNVIDPPIYARHRIIPMSWYGRITLASELLGCTEE 406
DB 2074 LFIQDGH--TKVFQGNDSSTPVNALDPLEFTRLRHPISMAOHIALREVLGCEAD 2131

RESULT 11
ID NRP2_RAT STANDARD: PRT: 925 AA.
AC 035276;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROPELIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).
OS NRP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=97433085; PubMed=9288754;
RA Kolodkin A.L., Levengood D.V., Rowe E.G., Tai Y.-T., Giger R.J.,
RA Ginty D.D.;
RT "Neuropilin is a semaphorin III receptor.";
RL Cell 90:753-762(1997).
CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3E, VEGF-165
CC -1- AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: FOUND IN CERTAIN NEURONAL POPULATIONS OF THE
CC CNS AND IN OTHER NONNEURONAL TISSUES INCLUDING MESENCHYMAL TISSUE
CC LINING IN THE RIBS.
CC -1- SIMILARITY: BELONGS TO THE NEUROPELIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF016297; AAC53338.1; -
CC InterPro: IPR000421; -
CC InterPro: IPR000859; -
CC InterPro: IPR000938; -
CC Pfam: PF00431; CUB; 2.
CC Pfam: PF00629; MAM; 1.
CC Pfam: PF00754; F5_F8_type_C; 2.
CC PROSITE: PS01180; CUB; 2.

```

```

DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 1.
DR PROSITE: PS50060; MAM_2; 1.
KW Transmembrane; Glycoprotein; Neuropein; Signal; Repeat; Receptor.
FT SIGNAL 1 22
FT CHAIN 23 925
FT DOMAIN 23 858
FT TRANSMEM 859 883
FT DOMAIN 884 925
FT DOMAIN 28 142
FT DOMAIN 149 267
FT DOMAIN 277 427
FT DOMAIN 434 592
FT DOMAIN 642 802
FT DISULFID 28 55
FT DISULFID 83 105
FT DISULFID 149 175
FT DISULFID 208 230
FT DISULFID 277 427
FT DISULFID 434 592
FT CARBOHYD 152 152
FT CARBOHYD 157 157
FT CARBOHYD 629 629
FT CARBOHYD 833 833
FT CARBOHYD 834 834
SQ SEQUENCE 925 AA; 103896 MW; 3BF62903F644851C CRC64;

Query Match 20.2%; Score 457; DB 1; Length 925;
Best Local Similarity 35.5%; Pred. No. 1,le-28;
Matches 120; Conservative 58; Mismatches 124; Indels 36; Gaps 15;

QY 81 YKSGSGPIEGGIIISNOOITASSRHALRGLOKQWYVYARLKKGLINMTAENDRKR 140
DB 275 FQCNAPLQMESGRITANQOISASST---ESDGRMTQOGRHL--GDNGMT-PVNSKKE 327
QY 141 WTQINLQKRMKRVGVITOGA-KRISPEYIKFYKLAYSNDGKTAMKVKGTNEDVYR 198
DB 328 YLOYDLRFLTMTLTAITOGAISRETQGYVYKYLEVSTNGEDMMVYR-HQKNH-KVFO 385
QY 199 GNIDNNTPYANSTPPIKQAYRLVPOVCRHCTLMELGCELSG--CSEPLGKSGHI 256
DB 386 ANNDTELVNKLKLTPLTRPIRPOYHGLTALTEFGCRVTDAPCSNMLGMSLI 445
QY 257 QDYQITASSIFRTLMMDFTPEPKARLDKQKQKVNAMTSGHNDOSQ---WLOYDLLVPT 312
DB 446 ADTOISASSTRE-----YLMSPSARL--VSSRSGWFP-RNPQAPQPEEMLYQVDLGTPK 496
QY 313 KVTGIIITOGAKDFGHV-----QFVGSYKLAYSNDGEHMTYODEKOKKDKVQGNFND 366
DB 497 TVKGVITOGAGGDSITAMCARAFVKKFKVSYSLNKKDMETIQDPRTOQPLFEGSMHYD 556
QY 367 THKKNVIDPPIYARHRIIP--MSWYGRITLASELLGC 402
DB 557 TPDIRRE-PVPAQYVYVPERMSPAG--TGMRLVGLGC 592

RESULT 12
NRP2_MOUSE STANDARD: PRT: 931 AA.
AC 035375; 035373; 035374; 035376; 035377; 035378;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROPELIN-2 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR 2).
OS NRP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

```


RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
RT Neupilin-1 is expressed by endothelial and tumor cells as an
RT isoform-specific receptor for vascular endothelial growth factor.*;
RL Cell 92:735-745(1998).

RN [3]
RP CHARACTERIZATION.

RA MEDLINE=20309748; PubMed=10748121;
RX Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;

RT Neupilin-2 and neupilin-1 are receptors for the 165-amino acid
RT form of vascular endothelial growth factor (VEGF) and of placenta
RT growth factor-2, but only neupilin-2 functions as a receptor for
RT the 145-amino acid form of VEGF.*;
RL J. Biol. Chem. 273:18040-18045(2000).

CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR SEMAPHORINS 3C, 3F, VEGF-165
CC AND VEGF-145 ISOFORMS OF VEGF, AND THE PLGF-2 ISOFORM OF PGF.
CC -1- SUBUNIT: NEUPILIN-2 PROBABLY FORMS A HETEROMERIC COMPLEX WITH
CC NEUPILIN-1 IN ORDER TO BE A FUNCTIONAL SEMAPHORIN 3C RECEPTOR.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A22 (SHOWN HERE), A0 AND A17;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- SIMILARITY: BELONGS TO THE NEUPILIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: AF02859; AAC51788.1; -;
DR EMBL: AF02860; AAC51789.1; -;
DR EMBL: AF016098; AAC12922.1; -;
DR MIM: 602070; -;

DR InterPro: IPR000421; -;
DR InterPro: IPR000855; -;
DR InterPro: IPR000958; -;

DR Pfam: PF00431; CUB; 2;
DR Pfam: PF00629; MAM; 1;
DR Pfam: PF00754; F5_F8_type_C; 2;

DR Prosite: PS01180; CUB; 2;
DR Prosite: PS01285; FA58C_1; 2;
DR Prosite: PS01286; FA58C_2; 2;
DR Prosite: PS50060; MAM; 2; 1.

KW Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
KW Alternative splicing.

FT SIGNAL 1 20 OR 22 (POTENTIAL).
FT CHAIN 21 931 NEUPILIN-2.
FT DOMAIN 21 864 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 865 869 POTENTIAL.
FT DOMAIN 931 931 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 142 CUB 1.
FT DOMAIN 149 267 CUB 2.

FT DOMAIN 277 427 F5/8 TYPE C 1.
FT DOMAIN 434 592 F5/8 TYPE C 2.
FT DOMAIN 642 802 MAM.
FT DOMAIN 671 674 POLY-SPR.
FT DISULFID 83 105 BY SIMILARITY.
FT DISULFID 149 175 BY SIMILARITY.
FT DISULFID 208 230 BY SIMILARITY.
FT DISULFID 277 427 BY SIMILARITY.
FT DISULFID 434 592 BY SIMILARITY.

FT CARBOHYD 152 152 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC...) (POTENTIAL).
FT VASPLIC 809 813 MISSING (IN ISOFORM A17).
FT VASPLIC 809 830 MISSING (IN ISOFORM A0).
FT CONFLICT 602 602 E -> K (IN REF. 1).

SO SEQUENCE 931 AA: 104830 MW: 270CBAE69A0A797C CRC64;

Query Match 20.0%; Score 453; DB 1; Length 931;
Best Local Similarity 35.4%; Pred. No. 2,2e-28;

Matches 120; Conservative 59; Mismatches 122; Indels 38; Gaps 15;

OY 81 YKSGPGICIGGIIISMOOTASTSTRALFGLOKVPYARLNKGLINAMTAENDRKR 140
DB 275 FQCVPLGMSGGINAEQISAST----YSDGRTPQOSLH--GDDNWT-PILDSNKE 327
OY 141 WIQINLORKNRVTGVITGCA--KRIGSEYIKFYKLAIVSDGKTAMKVKGTEDMYFR 198
DB 328 YLQADLEFLMTALTAITQGAISRETONGYKVSYLEVSTNGCDMMVYR-HCKNH-KVFQ 385
OY 199 GNIDNNTPYANSTFPPIKAQVRYLPQVCRHCHLIMELIGCELSG--CSEPLGKSGHI 256
DB 386 ANNDATEVVLNKLHAPLHFRVIRIPQTHWSGIALRLEFGCVTDAPCSNMLGMLGLI 445
OY 257 ODVOITASSIFRLNMDMTPEKARL--DKQS---KVAMTSGHNSQSMLOVDILVLP 311
DB 446 ADSQISASS-----TOEYLMSPSARLVSRSGWFPRIQAOPE---EEMLOVDLSTP 495
OY 312 TKVTGIIITGCAKDFGHV-----QFVSGYKLAVNSDGEHWTVODEKRRKRVQGFND 365
DB 496 KTVGAVITGARGDSITAVEARAFVRKFKVSYSLNCKMEXIODPTQOPKLFEGNNHY 555
OY 366 DTHRKNVIDPPIYARHRIILP--MSWYGRITLASELIGC 402
DB 556 DTPDIRRD-PIPAQVYVYRPERMSPAG-IQMRLVGLVC 592

RESULT 14

ID NRRL_HUMAN STANDARD: PRT: 923 AA.
AC 014786; 060461.
DT 30-MAY-2000 (Rel. 39, created)
DT 01-OCT-2000 (Rel. 40, last sequence update)
DT 01-OCT-2000 (Rel. 40, last annotation update)
DE NEUPILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
GN NRPI OR NRP OR VEGF165R.
OS Homo sapiens (Human).
OC Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM).
RX MEDLINE=97433084; PubMed=9280753;
RA He Z., Tessier-Lavigne M.;
RT "Neupilin is a receptor for the axonal chemorepellent semaphorin
RT III.";
RL Cell 90:739-751(1997).
RN [2]
RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM), AND SEQUENCE OF 22-39.
RC TISSUE-Breast;
RX MEDLINE=98180099; PubMed=9529250;
RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
RT "Neupilin-1 is expressed by endothelial and tumor cells as an
RT isoform-specific receptor for vascular endothelial growth factor.*";
RL Cell 92:735-745(1998).
RN [3]
RP SEQUENCE FROM N.A. (SOLUBLE/SNRP ISOFORM), AND SEQUENCE OF 22-31.
RC TISSUE-Prostatic adenocarcinoma;
RX MEDLINE=20183929; PubMed=10688880;
RA Gagnon M.L., Bielenberg D.R., Gechtman Z., Miao H.-Q., Takashima S.,
RA Soker S., Klagsbrun M.;
RT "Identification of a natural soluble neupilin-1 that binds vascular
RT endothelial growth factor: In vivo expression and antitumor
RT activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2573-2578(2000).
RN [4]
RP CHARACTERIZATION.

RX MEDLINE-20309748; PubMed-10748121;
 RA Glutamate-Politorak Z., Cohen T., Herzog Y., Neufeld G.;
 RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid
 RT form of vascular endothelial growth factor (VEGF) and of placenta
 RT growth factor-2, but only neuropilin-2 functions as a receptor for
 RT the 145-amino acid form of VEGF";
 RL J. Biol. Chem. 275:18040-18045(2000).
 CC -1- FUNCTION: THE MEMBRANE-BOUND ISOFORM IS A RECEPTOR INVOLVED IN THE
 CC DEVELOPMENT OF THE CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE
 CC FORMATION OF CERTAIN NEURONAL CIRCUITS AND IN ORGANOREGULATION
 CC OUTSIDE THE NERVOUS SYSTEM. IT MEDIATES THE CHEMOREPULSANT
 CC ACTIVITY OF SEMAPHORINS. IT BINDS TO SEMAPHORIN 3A, THE PLGF-2
 CC ISOFORM OF PGF, THE VEGF-165 ISOFORM OF VEGF AND VEGF-B.
 CC COEXPRESSION WITH KDR RESULTS IN INCREASED VEGF-165 BINDING TO KDR
 CC AS WELL AS INCREASED CHEMOTAXIS. IT MAY REGULATE VEGF-INDUCED
 CC ANGIOGENESIS.
 CC -1- FUNCTION: THE SOLUBLE/SNRP1 ISOFORM BINDS VEGF-165 AND APPEARS
 CC TO INHIBIT ITS BINDING TO CELLS. IT MAY ALSO INDUCE APOPTOSIS BY
 CC SEQUESTERING VEGF-165. MAY BIND AS WELL VARIOUS MEMBERS OF THE
 CC SEMAPHORIN FAMILY. ITS EXPRESSION HAS AN AVERSE EFFECT ON BLOOD
 CC VESSEL NUMBER AND INTEGRITY.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE SOLUBLE/SNRP1
 CC ISOFORM IS SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: MEMBRANE-BOUND (SHOWN HERE)
 CC AND SOLUBLE/SNRP1. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: THE EXPRESSION OF MEMBRANE-BOUND AND
 CC SOLUBLE/SNRP1 ISOFORMS DOES NOT SEEM TO OVERLAP. THE MEMBRANE-
 CC BOUND ISOFORM IS EXPRESSED BY THE BLOOD VESSELS OF DIFFERENT
 CC TISSUES. IN THE DEVELOPING EMBRYO IT IS FOUND PREDOMINANTLY IN THE
 CC NERVOUS SYSTEM. IN ADULT TISSUES, IT IS HIGHLY EXPRESSED IN HEART
 CC AND PLACENTA; MODERATELY IN LUNG, LIVER, SKELETAL MUSCLE, KIDNEY
 CC AND PANCREAS; AND LOW IN ADULT BRAIN. THE SOLUBLE/SNRP1 ISOFORM IS
 CC FOUND IN LIVER HEPATOCYTES, KIDNEY DISTAL AND PROXIMAL TUBULES.
 CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)

FT DOMAIN 645 811 MAM.
 FT DISULFID 27 54 PROBABLE.
 FT DISULFID 82 104 PROBABLE.
 FT DISULFID 147 173 PROBABLE.
 FT DISULFID 206 228 PROBABLE.
 FT DISULFID 275 424 BY SIMILARITY.
 FT DISULFID 431 583 BY SIMILARITY.
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 642 644 EPP -> GIK (IN SOLUBLE/SNRP1 ISOFORM).
 FT VARSPLIC 645 923 MISSING (IN SOLUBLE/SNRP1 ISOFORM).
 FT CONFLICT 26 26 K -> E (IN REF. 1).
 FT CONFLICT 749 749 D -> H (IN REF. 2).
 FT CONFLICT 855 855 E -> D (IN REF. 2).
 SQ SEQUENCE 923 AA; 103120 MW; ADADCA4A849E5057 CRC64;
 Query Match 19.7%; Score 446.5; DB 1; Length 923;
 Best Local Similarity 35.5%; Pred. No. 7.1e-28;
 Matches 119; Conservative 50; Mismatches 135; Indels 31; Gaps 12;
 QY 81 YKCSGPIEGEGITISNOGITASSTHRALTEGLQKMYRYVARLKKGLIYATAAENDMKR 140
 DB 273 FKMEALGHESELISDQITASSQIST-----NMSAKSRLLYPE--NGWTPE-DSYRE 324
 QY 141 WQIQLRKKRVYGYTQGA--KRIGSPEYIKFYKIAYSNDGKTWAMYKVKINEDVFR 198
 DB 325 MIVDVLGLRFTAVAGOGAISKETKKKKYVKKYKIDVSSNEDM--TTKEGKKPVLQ 382
 QY 199 GNDNNTPYANSFPPIKACYRLYPVCCRHCTLMELGCELSG--CSEPLMKSGHI 256
 DB 383 GNNPDPVYVAFPRPLTRVRIRKIPATMETGISMFEYVGKIDYCSGMLGWSGLI 442
 QY 257 QDYQTASSFTFLNMDKFTPEPKARL--DKQKKNATP---SCHNDOSQMLVDLAV 311
 DB 443 SDSQITSS-----NQGDRMMRPENILVYTSRSG---WALPPAHSTINMLQIDGEE 492
 QY 312 TKVGTITOGAKDFHQVFGVSYKLAYSNDGHWYVYODEKORCKDYKDFNDTHRN 371
 DB 493 KIVRGIILOGGHRENKVFMRKFKIGYSNNGSDMKIMDSKRAKSPFGNNYDFPELR 552
 QY 372 VIDPPIYARHRIILP--MSWYGRITLASELIGCTEE 405
 DB 553 TP-PALSTPFIYIPERRATHGGLRMELLCVE 586
 RESULT 15
 NRPL_MOUSE STANDARD; PRT; 923 AA.
 AC P97333;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN).
 GN NRPL OR NRE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Embryonic brain;
 RX MEDLINE-96353149; PubMed-8748368;
 RA Kawakami A., Kitsuikawa T., Takagi S., Fujisawa H.;
 RT "Developmentally regulated expression of a cell surface protein,
 RT neuropilin, in the mouse nervous system.";
 RL J. Neurobiol. 29:1-17(1996).
 CC -1- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
 CC CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
 CC CERTAIN NEURONAL CIRCUITS AND IN ORGANOREGULATION OUTSIDE THE NERVOUS

CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS. IT
 CC BINDS TO SEMAPHORIN 3A, THE PLGF-2 ISOFORM OF PGF, THE VEGF-165
 CC ISOFORM OF VEGF AND VEGF-B. COEXPRESSION WITH KDR RESULTS IN
 CC INCREASED VEGF-165 BINDING TO KDR AS WELL AS INCREASED CHEMOTAXIS.
 CC IT MAY REGULATE VEGF-INDUCED ANGIOGENESIS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM.
 CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: D50086; BAA08789.1; -
 DR MGD: MGI:106206; NRP.
 DR InterPro: IPR000421; -
 DR InterPro: IPR000859; -
 DR InterPro: IPR000998; -
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR Pfam: PF00629; MAM; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 DR PROSITE: PS00740; MAM_1; 1.
 DR PROSITE: PS50060; MAM_2; 1.
 DR KEGG: K01100; Glycophorin; Glycophorin.
 FT SIGNAL 1 21
 FT CHAIN 1 923
 FT DOMAIN 22 923 NEUROFILIN-1.
 FT TRANSMEM 22 836 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 857 879 POTENTIAL.
 FT DOMAIN 880 923 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 27 141 CUB 1.
 FT DOMAIN 147 265 CUB 2.
 FT DOMAIN 275 424 F5/8 TYPE C 1.
 FT DOMAIN 431 583 F5/8 TYPE C 2.
 FT DOMAIN 645 811 MAM.
 FT DISULFID 27 54 PROBABLE.
 FT DISULFID 82 104 PROBABLE.
 FT DISULFID 147 173 PROBABLE.
 FT DISULFID 206 228 PROBABLE.
 FT DISULFID 275 424 BY SIMILARITY.
 FT DISULFID 431 583 BY SIMILARITY.
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 923 AA; 103020 MM; 0644B8A170796808 CRC64;

Query Match 19 6%; Score 443.5; DB 1; Length 923;
 Best Local Similarity 35.8%; Pred No. 1.2e-27;
 Matches 120; Conservative 48; Mismatches 136; Indels 31; Gaps 12;

OY 81 YKSGPLGIEGGLISNQITASSSTRALFGLOKWPYARLNKKGLINAWTAENDRWKR 140
 DB 273 FKCMALGMESEGLHSQITASSOYGT-----NMSVERSRLNYPE--NCWTGCE--DSYKE 324
 OY 141 WTQINLQKRMVGTGVTGGA--KRISPEYIKFYKIAYSNDGKTWAMYKVGNTEDMVR 198
 DB 325 WTQVDLGLRVTAVGTGGAISKETKKYKVTYRVDISSNGEDW--ISLKEGNRAIIPQ 382
 OY .199 GNIDNNTPTANSFTPTKAYRYLPQVCRHCTRMELGCELSG--CSEPLGMSGSHI 256
 DB 383 GTNTPTDVLGVFSKPLITFRVRIKIPVSWETGISMRFEVIGCKLTIDYPCSGMLGWSGLI 442

OY 257 QDVQITASSIFFTLNNDFTWEPKARL--DKQKVNAMT---SGHNDQSQWLDVLLVP 311
 DB 443 SDSQITASS-----NQADRNMMPEIRILVTSRTG---NALPSPHPYTNEMLQVDLGE 492
 OY 312 TKVYGIITOGAKDFGVOPVGSYKLAYSNDGEHWTYVQDEKOKKVPQGNFNDTRKN 371
 DB 493 KIVAGYILGGKHREKNKVPMRKFKIAYSNNQSDPKITMDSKRKAKSPFGNNNTDTPELR 552
 OY 372 VIDPPIYARHRIPLP--WSWYGRITILASLLCTIDE 405
 DB 553 TFS-PLSTRFIRIYPERATVSGLRLMELGCEVE 586

Search completed: May 23, 2001, 06:19:00
 Job time: 364 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:20:46 ; Search time 105.14 Seconds

(without alignments)
452,600 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Sequence: 1 GPCTPMPCNGSTGCTEISFAY.....MSWGRITLASELCTGTFEE 406

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:
1: SP_archaea:
2: SP_bacteria:
3: SP_fungi:
4: SP_human:
5: SP_invertebrate:
6: SP_mammal:
7: SP_mhc:
8: SP_organelle:
9: SP_phage:
10: SP_plant:
11: SP_rodent:
12: SP_unclassified:
13: SP_vertebrate:
14: SP_virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2229.5	98.5	480	043854	043854 homo sapien
2	2180.5	96.3	480	035474	035474 mus musculu
3	1096	48.4	463	11 09R1X9	09R1X9 mus musculu
4	1079.5	47.7	426	11 09WTS3	09WTS3 mus musculu
5	933.5	41.2	363	6 077718	077718 equus caball
6	682.5	30.1	2183	11 088783	088783 mus musculu
7	670.5	29.6	2224	4 043737	043737 homo sapien
8	648	28.4	2343	6 018806	062730 canis faml
9	644	28.4	2343	6 018806	018806 canis faml
10	477	21.1	858	5 076470	076470 lytechinus
11	464.5	20.5	216	4 014286	014286 homo sapien
12	464.5	19.7	779	4 09MTT3	09MTT3 homo sapien
13	436.5	18.3	921	11 09OX38	09OX38 rattus norv
14	428.5	18.9	192	11 09XMO1	09XMO1 mus musculu
15	312.5	13.8	3644	5 09V094	09V094 drosophila
16	312.5	13.8	3644	5 09V050	09V050 drosophila
17	279	12.3	224	11 09R1M6	09R1M6 mus musculu
18	265.5	11.7	1128	11 088442	088442 mus musculu
19	262.5	11.6	1158	4 014113	014113 homo sapien

20	248.5	11.0	764	11 054860	054860 mus musculu
21	248	11.0	2352	5 061240	061240 halocynthia
22	237.5	10.5	752	13 042374	042374 brachydanio
23	234	10.3	734	4 09NMB5	09NMB5 homo sapien
24	232	10.2	2319	11 09R172	09R172 rattus norv
25	229	10.1	728	6 097567	097567 bos taurus
26	228	10.1	2531	5 016004	016004 lytechinus
27	227.5	10.0	2653	5 025253	025253 lucilia cup
28	225	9.9	719	11 061281	061281 mus musculu
29	224.5	9.9	156	5 026661	026661 strongyloce
30	224.5	9.9	2634	5 09W4T8	09W4T8 drosophila
31	224.5	9.9	2704	5 097458	097458 drosophila
32	223.5	9.9	728	13 090656	090656 gallus galli
33	223	9.8	2447	13 013149	013149 fugu rubrip
34	223	9.8	2471	11 09OW30	09OW30 rattus sp.
35	221.5	9.8	529	5 025058	025058 heliocidari
36	221	9.8	653	4 09N0T9	09N0T9 homo sapien
37	221	9.8	685	4 09NR61	09NR61 homo sapien
38	221	9.8	1203	11 006008	006008 mus musculu
39	221	9.8	2470	11 035516	035516 mus musculu
40	220.5	9.7	861	11 09OW58	09OW58 mus sp. mot
41	220	9.7	406	5 025059	025059 heliocidari
42	220	9.7	1530	11 09WUG5	09WUG5 rattus norv
43	220	9.7	1531	11 088279	088279 rattus norv
44	218.5	9.7	717	13 087357	087357 brachydanio
45	218.5	9.7	722	11 09Z100	09Z100 mus musculu

ALIGNMENTS

RESULT	1	ALIGNMENTS
ID	043854	PRELIMINARY: PRT; 480 AA.
AC	043854; 043855;	
DT	01-JUN-1998 (TREMBLrel. 06, Created)	
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)	
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	
DE	INTEGRIN-BINDING PROTEIN DELI PRECURSOR.	
GN	DELI.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX	NCBI_Taxid=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.	
RC	TISSUE-EMBRYONIC LUNG;	
RC	MDLINE=98083109; PubMed=9420328;	
RA	Hidal C., Zupancic T.J., Penta R., Mikhail A., Kawana M.,	
RA	Quertermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,	
RA	Auerbach R., Hogan B.L.M., Snodgrass R., Quertermous T.;	
RT	"Cloning and characterization of developmental endothelial locus-1: an	
RT	embryonic endothelial cell protein that binds the alphabeta3 integrin	
RT	receptor.";	
RL	Gene: Dev. 12:21-33(1998).	
CC	-1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH	
CC	INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS	
CC	FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN	
CC	REGULATION OF VASCULAR MORPHOGENESIS OF REMODELING IN EMBRYONIC	
CC	DEVELOPMENT.	
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.	
CC	-1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND	
CC	SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.	
CC	-1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.	
CC	-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.	
DR	EMBL: U70312; AAC02648.1; -;	
DR	EMBL: U70313; AAC02649.1; -;	
DR	HSSP: P00740; IIXA.	
DR	INTERPRO: IPR000152; -;	
DR	INTERPRO: IPR000421; -;	
DR	INTERPRO: IPR000561; -;	
DR	INTERPRO: IPR000742; -;	
DR	INTERPRO: IPR001881; -;	

DR PFAM; PF00008; EGF; 3.
 DR PFAM; PF00754; F5_F8_type.C; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR EGF-like domain; Alternative splicing; Signal; Developmental protein;
 KM Cell adhesion; Repeat; Vascularization.
 FT SIGNAL 1 16
 FT CHAIN 17 480
 FT DOMAIN 26 59
 FT DOMAIN 78 116
 FT DOMAIN 123 154
 FT DOMAIN 161 311
 FT DOMAIN 322 473
 FT SITE 96 98
 FT DISULFID 26 37
 FT DISULFID 31 48
 FT DISULFID 50 59
 FT DISULFID 78 89
 FT DISULFID 83 105
 FT DISULFID 107 116
 FT DISULFID 158 314
 FT DISULFID 301 305
 FT DISULFID 319 476
 FT VARSPLIC 66 66
 FT VARSPLIC 67 76
 SQ SEQUENCE 480 AA; 53765 MW; F7171E23A309FD48 CRC64;

Query Match 98.5%; Score 2229.5; DB 4; Length 480;
 Best Local Similarity 99.0%; Pred. No. 6,7e-192;
 Matches 402; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 1 GPCPPHNGHNGCTSEARQPTPLGYCKCPRGFNCIHQHNNECEYERKNGGCT 60
 DB 76 GPCPPHNGHNGCTSEARQPTPLGYCKCPRGFNCIHQHNNECEYERKNGGCT 135
 OY 61 DLVANYSECEPGEFMRNCQYKCSPLGEGGISNQDITASSTRALFGLQKWPYYAR 120
 DB 136 DLVANYSECEPGEFMRNCQYKCSPLGEGGISNQDITASSTRALFGLQKWPYYAR 195
 OY 121 LNKKGILNMTAAENDRKFMIDINQRMKRVGTITOGAKRISPEYIKFYIAVNSNG 180
 DB 196 LNKKGILNMTAAENDRKFMIDINQRMKRVGTITOGAKRISPEYIKFYIAVNSNG 254
 OY 181 KTMANYKKKGTNEDMVFRTGNDNNTPYANSFPPIKAQYVLYPOVCRRHCTLMELLGC 240
 DB 255 KTMANYKKKGTNEDMVFRTGNDNNTPYANSFPPIKAQYVLYPOVCRRHCTLMELLGC 314
 OY 241 ELSCGSEPLGKSGHIDYQITASSIFRTLNMDFTWEPKARLDKOGKVNATSGHND 300
 DB 315 ELSCGSEPLGKSGHIDYQITASSIFRTLNMDFTWEPKARLDKOGKVNATSGHND 374
 OY 301 SOMLOVDLAVTKYTGITITOGANDFGHVOFVSXYKLAVSNGEMWYIYOEDKQRKQVQ 360
 DB 375 SOMLOVDLAVTKYTGITITOGANDFGHVOFVSXYKLAVSNGEMWYIYOEDKQRKQVQ 434
 OY 361 GNFDMDTRKKNVDIPPIYARHRIIPWSWYGRITLASELGCTEE 406
 DB 435 GNFDMDTRKKNVDIPPIYARHRIIPWSWYGRITLASELGCTEE 480

RESULT 2
 ID 035474 PRELIMINARY; PRT: 480 AA.
 AC 035474; 035475;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.

GN EDL3 OR DEL1.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NX NCBI TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
 RC TISSUE=EMBRYO;
 RX MEDLINE=98083109; PubMed=9420328;
 RA Hidaï C., Zupancic T., Penta K., Mikhail A., Kawana M.,
 RA Quettemous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
 RA Auerbach R., Hogan B.L.M., Snodgrass R., Quettemous T.; an
 RA "Cloning and characterization of developmental endothelial locus-1: an
 RA embryonic endothelial cell protein that binds the alphavbeta3 Integrin
 RA receptor.";
 RL Genes Dev. 12:21-33(1998).
 CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
 CC INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
 CC FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
 CC REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC
 CC DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND
 CC SHORT ISOFORM. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
 CC CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP
 CC OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
 CC NEURONS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7, AFTER DAY
 CC 15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME
 CC OF BIRTH.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 DR EMBL: AF031524; AAB86585.1; -.
 DR EMBL: AF031525; AAB86586.1; -.
 DR HSSP: P00740; 1EDM.
 DR MGD: MGI:1329025; Edl13.
 DR INTERPRO: IPR000152; -.
 DR INTERPRO: IPR000421; -.
 DR INTERPRO: IPR000561; -.
 DR INTERPRO: IPR000742; -.
 DR INTERPRO: IPR001881; -.
 DR PFAM: PF00008; EGF; 3.
 DR PFAM: PF00754; F5_F8_type.C; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01187; EGF_CA; 1.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 DR EGF-like domain; Alternative splicing; Signal; Developmental protein;
 KW Cell adhesion; Repeat; Vascularization.
 FT SIGNAL 1 16
 FT CHAIN 17 480
 FT DOMAIN 26 59
 FT DOMAIN 78 116
 FT DOMAIN 123 154
 FT DOMAIN 161 311
 FT DOMAIN 322 473
 FT SITE 96 98
 FT DISULFID 26 37
 FT DISULFID 31 48
 FT DISULFID 50 59
 FT DISULFID 78 89
 FT DISULFID 83 105
 FT DISULFID 107 116
 FT DISULFID 158 314
 FT DISULFID 301 305
 FT DISULFID 319 476
 FT VARSPLIC 216 221
 FT VARSPLIC 222 480
 SQ SEQUENCE 480 AA; 53740 MW; 4CD91EF9261714D CRC64;


```

Db 28 CDSCLNGTCLTGDNNDY-----CLCEPFTGLVC---NTERGCPSPNCPY 74
QY 55 NSGICT-----DLVANSCECPGEFMGRNCOYKCGSPGLIEGIIISNOQITASTHR 106
Db 75 NDAKCLVTLDPQRGDIFTEYICCPGVYSGICHCEGSCQGLMGEGAIASQISASSVYM 134
QY 107 ALFGLOKWPYARLNNKGLINAMTAENDRMKRWIOINLORKRNVGILTOGAKRIGSP 166
Db 135 GPMGLQRMWGPRLARLYRTGIVNMTASNDY- KPMIOYNLLRKKRNVGVTQASAGRA 193
QY 167 EIKRKYLAISNDGKTAMKVKGTMEDEVFRGNIDNTPYANSETPPIKARLVPRXY 226
Db 194 ELKFTFKVAISLDGRKFEFTIDE- SGGDEKFEIGLNDNNSLKVNMFPNLTLOQIRLPPVS 252
QY 227 CRRHCTLMELLGCELSGCEPLGKMSGHIIDYQITASSIFRTLMDMFTWBRKALDK 286
Db 253 CHRGCTLFFELLGCELGSEPLGLKNNTIPDSOMASSSYKTNLRAFGWYFHLRLDN 312
QY 287 QCKVAMTSGHNDOSQWLQYDLVLPKVTGILTOGAKDFGHVQFVGSYKLAYSNDGEHMT 346
Db 313 QCKINAMTAQNSAKEMWLOYDLGTQORVGTITGARDGFHIQYVASYKVAHSDGCVMT 372
QY 347 VYQDEKOKKDKVPOGNDNTHRKNVIDPEIARHRIILPMSWGRITLASSELLGC 402
Db 373 VI--EQOSSKVFQGLNDNNSHKNIFEKFMARIVRLVPSWNRILTLLELLGC 426

RESULT 5
ID 07718 PRELIMINARY; PRT; 363 AA.
AC 07718;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, last sequence update)
DE 01-MAY-2000 (TREMblrel. 13, last annotation update)
DE SPEEM-MEMBRANE ASSOCIATED PROTEIN P47 (FRAGMENT).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_Taxid=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Genzel M., Toepfer-Petersen E.;
RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ010121; CA009010.1; -.
DR INTERPRO: IPR000421; -.
DR PRAM: PR00008; EGF_1;
DR PROSITE: PS00022; EGF_1; UNKNOW_2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FA58C_1; 1.
KW Sperm.
FT NON_TER 1
FT NON_TER 363
SQ SEQUENCE 363 AA; 40744 MW; 1F8B6395AF32338D CRC64;

```

Query Match 41.2%; Score 933.5; DB 6; Length 363;
 Best Local Similarity 44.2%; Pred. No. 9,1e-76;
 Matches 175; Conservative 63; Mismatches 95; Indels 63; Gaps 5;

```

QY 1 GPCITNPGHNGTCE-ISEAYRGDTFTGYVCKCPRGFNHICQHININECEVEPCCKNGIC 59
Db 30 GCFENPQNDGCHVIDSHRGDVFTOYICSPRGYGTGCEYT-----74
QY 60 TDLVANSCECPGEFMGRNCOYKCGSPGLIEGIIISNOQITASTHRALGLOKWPYAR 119
Db 75 -----CAMPLEMTGALADAOISASSVYGFPMGLQRMWPELA 111
QY 120 BLNKGGLINAMTAENDRMKRWIOINLORKRNVGILTOGAKRIGSEYIKFKYKTIASND 179
Db 112 RHFGTGIYVAMTASNDYK-NPMIOVALMKRKNVGTIVQASRGSTAEYIKTKYVAYSVD 170

```

```

QY 180 KXTAMKVKVKTNEDEVFRGNIDNTPYANSETPPIKAQVRLVPOYCRHRCTLMELLG 239
Db 171 GRKFOFIRDAGDSKDYFVGNVNDNSGLKVNMFEDVLETTYVRLVPLACHH-----220
QY 240 CELSGCEPLGKMSGHIIDYQITASSIFRTLMDMFTWBRKARLDKQCKVAMTSGHND 299
Db 221 ----GCTRP-----DRQITASTYRTGILNAFWSYEFYARLDKQCKFMANMTAQNS 267
QY 300 OSQWLQYDLVLPKVTGILTOGAKDFGHVQFVGSYKLAYSNDGEHMTVYQDEKOKKDYF 359
Db 268 ASEMLODDLSQKVEYVGTIGAKRDFGHIOYVDATKVSNSHDGAMNTEYRDQAADSKIF 327
QY 360 QGNFNDTHRKNVIDPEIARHRIILPMSWGRITL 395
Db 328 LGNDNNSHKNMFETPFLAREVRLPVAWHKRITL 363

RESULT 6
ID 088783 PRELIMINARY; PRT; 2183 AA.
AC 088783;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, last annotation update)
DE MURINE COAGULATION FACTOR V.
GN F5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang T.L., Cui J., Rehmuntulla A., Moussalli M., Kaufman R.J..
RA Ginsburg D.;
RT "The structure and function of murine factor V and its inactivation by
RT protein C.";
RL Blood 91:0-0(0012).
RN [2]
RP SEQUENCE FROM N.A.
RA Ginsburg D., Yang T.L., Cui J., Yang A.;
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U52925; AAC9553.1; -.
DR HSSP: P00450; 1KCM.
DR MGD: MGI:88382; F5.
DR INTERPRO: IPR000421; -.
DR PRAM: PR000394; Cu-oxidase; 3.
DR PROSITE: PS00754; F5_F8_Type_C; 2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
SQ SEQUENCE 2183 AA; 247228 MW; BF0A8AA723F60317 CRC64;

```

Query Match 30.1%; Score 682.5; DB 11; Length 2183;
 Best Local Similarity 43.2%; Pred. No. 3,4e-52;
 Matches 140; Conservative 51; Mismatches 122; Indels 11; Gaps 3;

```

QY 82 KCSGPELIEGIIISNOQITASTHRALGLOKWPYARLNNKGLINAMT---AAENDRW 138
Db 1865 ECKMPGSLGTVSDSQIRASEY-----LTYWRLARLNNKSGSTNAISIKRTALDPEI 1918
QY 139 KRWIOINLORKRNVGILTOGAKRIGSPYIKFYKIAYSNDGRTWAMYVKGTNEDWYFR 198
Db 1919 KPWIOVDMQKEVYVVTGIGIOGAKHYLKSCFTTERFOVAVSSDQTNMOIFPGKSGKSVMTFT 1978
QY 199 GNINNNPYANSETPPIKAQVRLVPOYCRHRCTLMELLGCELSGCEPLGKMSGHIID 258
Db 1979 GNSDGSITKENRLDPEIYARIRIHPTKSYNRPRLRLGCGCVNCSPLGLGEGRIOD 2038
QY 259 YQITASTIFRLNMDMFTWBRKARLDKQCKVAMTSGHNDOSQWLQYDLVLPKVTGTGII 318

```

Db 2039 KQITASSFKKSMWGDY--WEPSEIARLNAOGRVANAMQAKANNKQWLOVDLTKIKVTAIV 2096
 QY 319 TOGAKDEGHQFVGSYKLAISNGEHWITYQDEKQKDKVFGQNFNDTHRKVNDIPPIY 378
 Db 2097 TOGCKSLSSSMYKYSYIOISDGVAMKPYROKSMVDKIFEGNSNTKGMKNFNPPII 2156
 QY 379 ARHRIILPMSWGRITLASSELLGC 402
 Db 2157 SRFRIIRIPKWTNOSIALRLELFGC 2180

RESULT 7
 ID 043737 PRELIMINARY; PRT: 2224 AA.
 AC 043737;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE FACTOR V.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bird C.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: 299572; CAB6748.1; .
 DR HSSP: P00450; 1KCM.
 DR INTERPRO: IPR000421; .
 DR INTERPRO: IPR001117; .
 DR PFAM: PF00394; Cu-oxidase; 3.
 DR PFAM: PF00754; F5_F8_type_C; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 KW Blood coagulation.
 SQ SEQUENCE 2224 AA; 251673 MW; 1013108D49690EAB CRC64;

Query Match 29.6%; Score 670.5; DB 4; Length 2224;
 Best Local Similarity 41.8%; Pred. No. 4,1e-51;
 Matches 141; Conservative 54; Mismatches 127; Indels 15; Gaps 4;
 QY 69 ECGEFGNRCNOYKCSGPIEGIGISNQOITASSTHRALFGIOKWPYARLKNKGLIN 128
 Db 1897 QRPFLIMDDCRM---PGLSTGIIISDSQIKASEF-----LGTYEPRLALNNQSGTIN 1946
 QY 129 AMTA--AENDRKRWIQLNQRKMRVGTITOGAKRIGSPYIKFYKLAISNDGKTWAM 185
 Db 1947 AMSVEKLAIEFASKPWIQVDMQKEVITITIGTQGAHNYLKSCTTTEFYVAASSNQINMQI 2006
 QY 186 YKKKGTEENMVRFGNIDNNTPRANSTPRIKAOYVRLYQVCRHCTLMELLGCELSC 245
 Db 2007 FFGNSRNVWYFENGSDASTIKENQDPPIVARYIRISTPRAYNRPTLLELQGEVNC 2066
 QY 246 SEBLGKSGHIDYQITASSIFETILNMDFTWEPKARLDKQGVNAMSNDOSQMLQ 305
 Db 2067 STPLGEMNKIENKQITASSFKKSMWGDY--WEPFARLNAOGRVANAMQAKANNKQWLE 2124
 QY 306 VDLVETKVTGIIITOGAKDFGHVQFVGSYKLAISNGEHWITYQDEKQKDKVFGQNFND 365
 Db 2125 IDLTKIKITAITITOGCKSLSSSMYKYSYIYSEGQVEMKPYRLKSSWVDKIFEGNNT 2184
 QY 366 DTHRKVNDIPPIYARHRIILPMSWGRITLASSELLGC 402
 Db 2185 KGHVKNFNPPIISRFIRIVPXTWOSIALRLELFGC 2221

RESULT 8
 ID 062730 PRELIMINARY; PRT: 2343 AA.
 AC 062730;

DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE FACTOR VIII.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY, SPLEEN;
 RA Gordy P.W., Bowen R.A.;
 RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF049489; AAC05384.1; .
 DR HSSP: P00451; 1CEG.
 DR INTERPRO: IPR000421; .
 DR INTERPRO: IPR001117; .
 DR PFAM: PF00394; Cu-oxidase; 3.
 DR PFAM: PF00754; F5_F8_type_C; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADADD99 CRC64;

Query Match 28.6%; Score 648; DB 6; Length 2343;
 Best Local Similarity 42.6%; Pred. No. 4.6e-49;
 Matches 139; Conservative 53; Mismatches 118; Indels 16; Gaps 5;
 QY 82 KCSGPIEGIGISNQOITASSTHRALFGIOKWPYARLKNKGLINAMTAENDRKRW 144
 Db 2031 KCQPTLGMAASHIRDFQITASGQY-----QGMFKLARLHSSINAMSKD---FSW 2081
 QY 142 IQINLQRMKRVGTITOGAKRIGSPYIKFYKLAISNDGKTWAMYKKGTEENMVRFGNI 201
 Db 2082 IKVDLAMIITHGITOGAROKFSSLYVSPFIWYSLDGNKMHYSRGNSTGLWFFGNV 2141
 QY 202 DNNFPAVSFTPRPIKAQVRLYQVCRHCTLMELLGCELSCGSEPLGKMSGHIOYQI 261
 Db 2142 DSSGIRKNIENPPIIAQYIRLHPHYSTRITLREMLJGCDNSCMPLGMSKAIISDAQI 2201
 QY 262 TASSIFRTLNMDFTWEPKARLDKQGVNAMSNDOSQMLQVLDLVPKYVGTITOG 321
 Db 2202 TASSYLSM--LATWPSQARLHLQGTNANMRQANNPKMQLVDRKTKMYKVGITOG 2258
 QY 322 AKDEGHVQFVGSYKLAISNGEHWITYQDEKQKDKVFGQNFNDTHRKVNDIPPIYAR 380
 Db 2259 VKSLISMYKEFLISSQDHNMTLFLQNDKV--KVFQGNRDSSTPVNMLEPPIYAR 2315
 QY 381 HRIILPMSWGRITLASSELLGCTEE 406
 Db 2316 YVRLHPOSMAHHIALRLEVLGCDTQO 2341

RESULT 9
 ID 018806 PRELIMINARY; PRT: 2343 AA.
 AC 018806;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE FACTOR VIII.
 GN F8
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER.
 RA Cameron C., Nottley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
 RA Giles A., Lallier D.;
 RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.


```

DR INTERPRO, IPR001092; -.
DR PFAM, PF00431; CUB; 2.
DR PFAM, PF00629; MAM; 1.
DR PFAM, PF00754; F5_P8_type_C; 2.
DR PRINTS; PR000020; MAMDOMAIN.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS50060; MAM_2; 1.
SQ SEQUENCE 921 AA; 103052 MW; 58B29A9AA4978971 CRC64;

Query Match 19.3%; Score 436.5; DB 11; Length 921;
Best Local Similarity 35.5%; Pred. No. 1.3e-30;
Matches 119; Conservative 49; Mismatches 136; Indels 31; Gaps 12.

Qy 81 YKCSGFLGEGGIIISNQQTASTSTRHALGLOKWPYPYARLKKGLINMFWTAENDRMKR 140
Db 272 FRCMEALGEGSEIHSDQITASSQYGT-----NMSVERSLINPE--NGMTPEE-DSTYE 323
Qy 141 WQIINLQRMARYGTAVTQGA--KRISPEYIKFYKTAYSNDGKTWAMYVVKGTNEDMVR 198
Db 324 WQVVDGLGLRFTAVAGTQGAISKETKKYVYKTAIRDISNGEDW--ILKSGNKAILIQ 381
Qy 199 GNIDNNTPYANSFTPIKAQVYRLYPQVCRRHCTLMELIGCELS--CSEPLGKSGHI 256

```

DB	382	GNTNPTDVFVGFPKPLITTRREVRILPASMETGIMSRFEVYCKITDPYPCSGMLQVSLI	441
QY	257	QDVYTASSIPFTLNMDFWTEPRKARL--DKQCKVNAWT---SGHNDQSGMLQVLDLVP	311
Db	442	SSQSLTAS-----NQGRNNKMPENIRLVTISRTG---WALPSPHPYINEMLQVLDLGD	491
QY	312	TVTGTIIITGAKDEGHVGVFVSGSYKLYASNDSEHNTVYQDEKQSDKAFQCGNFNDTHRNK	371
Db	492	KIVRGVILIQGGHRRNKVFEMRKPAIYASNNQSDMKMLIMDSKKRAKSPFGNNNDPPELR	551
QY	372	VIDPPIYARHRIPL-WSWYGRITLASLCLCTEE	405
Db	552	AF-TELSTREIRIYPERATHSGILRMELLCGEVE	585
RESULT	14		
Q9QWQ1			
ID	Q9QWQ1	PRELIMINARY;	PRT; 192 AA.
AC	Q9QWQ1;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)		
DE	MILK FAT GLOBULE GLYCOPROTEIN (FRAGMENT).		

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
RT "Lactation-dependent expression of an mRNA splice variant with an exon
RT for a multiply O-glycosylated domain of mouse milk fat globule
RT glycoprotein Mfg-E8.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB021648; BAA36214.1; -;
DR EMBL; AB021647; BAA36214.1; JOINED.
DR INTERPRO; IPR000421; -;
DR INTERPRO; IPR000561; -;
DR INTERPRO; IPR001092; -;
DR PFMAM; PFM00008; EGF_1.
DR PFMAM; PFM00754; F5_F8_Type_C_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01285; FAF5C_1; 1.
FT NON_TER 1 1
FT NON_TER 192 192

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:15:15 ; Search time 75.05 Seconds
(without alignments)
309.237 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Perfect score: 2264
Sequence: 1 GPCPNPCNHGTCFCEISEAV.....MSWGRITLASELLGCTEE 406

Scoring table:

BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

A_Geneseq_0401:*

- 1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2264	100.0	513	20 W94685	Human Del-1 protei
2	2253	99.5	481	18 W10365	Human developmenta
3	2170.5	95.9	480	18 W10364	Mouse developmenta
4	2170.5	95.9	480	20 W94683	Murine Del-1 prote
5	1717	75.8	321	20 W94697	Human milk fat glo
6	1061.5	46.9	426	21 W94454	Mouse lactadherin
7	1020.5	45.1	387	16 R77252	HMFG 46 kDa antige
8	1020.5	45.1	387	21 Y94453	Human lactadherin
9	948	34.9	320	20 W98891	Peptide S60 ID NO:
10	790.5	34.9	221	18 W10366	Murine Del-1 trunc
11	784.5	34.7	221	20 W94684	Truncated murine D

12	682.5	30.1	2183	21 B03533	Murine factor V SE
13	675.5	29.8	2224	17 W04254	Human Factor V. H
14	670.5	29.6	2224	20 Y49564	Human lipoprotein
15	670	29.6	2319	19 W53485	Murine factor VIII
16	670	29.6	2319	19 W44135	Mus musculus facto
17	670	29.6	2319	20 Y31596	Mouse factor VIII
18	670	29.6	2319	22 B50467	Mouse factor VIII
19	655.5	29.0	2304	21 Y57848	Mouse Factor VIII
20	651	28.8	2351	7 P60741	Sequence of human
21	650	28.7	2351	8 P70448	Human Factor VIII:
22	650	28.7	2351	16 R78223	Human Factor-VIII:
23	649	28.7	684	16 R73022	Human Factor-VIII
24	649	28.7	684	16 R74091	Factor-VIII light
25	649	28.7	1014	8 P71139	Factor VIII:c varia
26	649	28.7	1383	18 W33227	Procoagulant-activ
27	649	28.7	1383	18 W33228	Procoagulant-activ
28	649	28.7	1383	18 W33229	Procoagulant-activ
29	649	28.7	1424	9 P80268	Modified factor VI
30	649	28.7	1424	10 P91169	Sequence of 740 Ar
31	649	28.7	1424	22 B48842	Mutant mature huma
32	649	28.7	1425	9 P80267	Modified factor VI
33	649	28.7	1438	21 B01262	B-domain deleted f
34	649	28.7	1440	12 R12971	Factor VIII:50.
35	649	28.7	1457	19 W46246	Human Factor VIII
36	649	28.7	1457	19 W44372	Human Factor VIII
37	649	28.7	1457	20 Y21675	Beta-domain delete
38	649	28.7	1471	18 W23414	Human B-domain del
39	649	28.7	1516	9 P80265	Modified factor VI
40	649	28.7	1661	18 W18670	Factor VIII-dB695-
41	649	28.7	2332	8 P71726	Factor VIII:c varia
42	649	28.7	2332	8 P71727	Factor VIII:c varia
43	649	28.7	2332	8 P71728	Factor VIII:c varia
44	649	28.7	2332	8 P71729	Factor VIII:c varia
45	649	28.7	2332	14 R43257	Human Factor VIII.

ALIGNMENTS

RESULT 1	
W94685	
ID	W94685 standard; Protein: 513 AA.
XX	
AC	W94685;
XX	
DT	05-MAY-1999 (first entry)
XX	
DE	Human Del-1 protein.
XX	
DE	Del-1: developmentally-regulated endothelial cell locus 1; cancer;
KW	discoilin 1; factor VIII-like domain; epidermal growth factor; EGF;
KW	diabetic retinopathy; rheumatoid arthritis; endometriosis;
KW	angiogenesis.
XX	
OS	Homo sapiens.
XX	
PN	US5877281-A.
XX	
PD	02-MAR-1999.
XX	
PF	05-JUN-1996; 96US-0659235.
XX	
PR	05-JUN-1996; 96US-0659235.
XX	
PA	(PROG-) PROGENITOR INC.
XX	(UYVA-) UNIV VANDERBILT.
PI	Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX	WPI: 1999-189720/16.
DR	N-PSDB; X18508.
XX	

Del-1 polypeptide sequences - useful for treatment of cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis

Claim 5; Fig 4; 73pp: English.

The present sequence is human developmentally-regulated endothelial cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1 proteins have an inhibitory effect on angiogenesis (blood vessel growth), this activity may be useful clinically to prevent neovascularisation of tissues such as tumour nodules and prevention of metastases. The anti-angiogenic activity of Del-1 may be used to treat abnormal conditions that result from angiogenesis, including cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be used to treat cardiac ischaemia, thrombotic stroke, wound healing and peripheral vascular disease. Del-1 is also useful for promoting bone formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis inducer.

Sequence 513 AA:

Query Match 100.0%; Score 2264; DB 20; Length 513;
Best Local Similarity 100.0%; Pred. No. 5,7e-148;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTPNPCNHNGTCEISEAYRQDFTFGYVCKCPRGFNGICHOHINCEVEPCKNGSICT 60
DB 108 gpcfpnpchngtgciseayrgdtflgyvckcprgfhgqhnlnccevepckngsict 167
QY 61 DLVANYSCCEPGFEMGRNCOYKCSGPILEGGIIISNOQITASSHRALEFLOKWPYYAR 120
DB 168 dlvanyscepgfemgrncqkcsplgleggilsnqgitasshraleflgkwkwyppar 227
QY 121 LNKGLINAWTAENDRMKRWIOINLQKRMRTGYITGAKRISPEYIKFYKANSNDG 180
DB 228 lnnkglinaawtaendrmkrtwqlnqtkmrvtyglitgaktlgspeyikfykaysndg 287
QY 181 KTNAMRYKVGKSTEDMYFRGNIDNTPYANSFPYKAQYVRLYQOVCRHCTLMELGCG 240
DB 288 ktnamrykvkstedyfrgnidntrpyansfpykaqyvrlyqovcrhctlmelg 347
QY 241 ELSCSEPLKMSGHIODVOITASSIFRTLNMDMTEPRKARLDKQKXVAMTSGHNDQ 300
DB 348 elscseplkmsghiodvoitassifrtlnmdmteprkarldkqkxvawtsgndg 407
QY 301 SQWLQVLDLVPTKVGITITQGAKEFGHVGQVGSYKLAYSNDGEMHTYVODEKORDKVPQ 360
DB 408 sqwlqvldlvptkvgititqgakdfghvgfvgysklaysndgemhtvygdekqrkdkvfg 467
QY 361 GNPNDRHNRKNIDPPYARHRIILPMSWYGRITLASELGLCTEE 406
DB 468 gnfnldrhnrvnidppiyarhriilpmswygritlaseelgctee 513

RESULT 2

W10365 W10365 standard; Protein: 481 AA.

AC W10365;

XX 03-MAY-1997 (first entry)

XX Human developmentally-regulated endothelial cell locus-1 protein.

XX Del-1; developmentally-regulated endothelial cell locus-1;
XX signal transduction; cancer; tumour marker; angiogenesis;
XX diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX cardiac ischaemia; stroke; vascular disease; wound healing;
XX vulnereary; bone formation; diagnosis; therapy.

OS Homo sapiens.
XX

FT Key Location/Qualifiers
FT Peptide 3...23
FT /label= Sig_peptide
FT Protein 24..481
FT /label= Mat_protein
FT Domain 26..59
FT /label= EGF-1
FT /note= "epidermal growth factor-1like domain 1"
FT Domain 78..116
FT /label= EGF-2
FT /note= "epidermal growth factor-1like domain 2"
FT Domain 123..154
FT /label= EGF-3
FT /note= "epidermal growth factor-1like domain 3"
FT Domain 158..315
FT /label= discoidin-1
FT /note= "discoidin I/factor VIII-like domain 1"
FT Domain 320..477
FT /label= discoidin-2
FT /note= "discoidin I/factor VIII-like domain 2"

MO9640769-A1.

PD 19-DEC-1996.

PP 05-JUN-1996; 96MO-US09456.

PR 07-JUN-1995; 95US-0480229.

PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.

PI Hogan B; Quertermous T, Snodgrass HR, Zupancic TJ;

DR WPI, 1997-052233/05.

DR N-PSDB: T47343.

PT New developmentally regulated endothelial cell locus-1 (del-1) gene
PT - used to develop prods. for the diagnosis and treatment of cancer
PT and conditions involving abnormal angiogenesis

PS Claim 4; Fig 6; 137pp: English.

CC Human Del-1 (W10365) is the polypeptide product of the human
CC developmentally-regulated endothelial cell locus-1 (del-1) gene
CC (T47343). It shows 94% amino acid homology to the mouse Del-1
CC protein (W10364). Structurally, members of this novel gene family
CC contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
CC domains. Del-1 is expressed in endothelial and certain tumour
CC cells. Its ability to inhibit vascular formation allows its used
CC as an anti-angiogenic agent. It can be used as a tumour marker,
CC to identify Del-1 binding partners, and to modulate endothelial
CC cell growth and blood vessel formation. Recombinant Del-1 can be
CC produced in transformed host cells utilizing vectors incorporating
CC del-1 nucleic acids.

SQ Sequence 481 AA:

Query Match 99.5%; Score 2253; DB 18; Length 481;
Best Local Similarity 99.5%; Pred. No. 3e-147;
Matches 404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCTPNPCNHNGTCEISEAYRQDFTFGYVCKCPRGFNGICHOHINCEVEPCKNGSICT 60
DB 76 gpcfpnpchngtgciseayrgdtflgyvckcprgfhgqhnlnccevepckngsict 135
QY 61 DLVANYSCCEPGFEMGRNCOYKCSGPILEGGIIISNOQITASSHRALEFLOKWPYYAR 120
DB 136 dlvanyscepgfemgrncqkcsplgleggilsnqgitasshraleflgkwkwyppar 195
QY 121 LNKGLINAWTAENDRMKRWIOINLQKRMRTGYITGAKRISPEYIKFYKANSNDG 180

Db 196 LnkKjlinawlaeandrnwrlqinlqrkmrvtgltgskrlyspkfyklynsndg 255
 Qy 181 KTWAMYKVKGTNDMEVFRGNIDNNTPYANSFPPKKAQYVLYQVCRHCTRMELTGC 240
 Db 256 ktwamykvkgtneawfrgnidnntpyansfppkkaqyvllyqvcrrhctrmel1gc 315
 Qy 241 ELSCSEPLGMSKSGHIDYQITASSIFRLNMDMTWEPKRAKLDKQKYNAMTSGHNDQ 300
 Db 316 elsgcseplgmksghidqyqitassifrlmdmtwepkrakldkqgkynawtsgndq 375
 Qy 301 SOWLOYDLVPTKVTGIIITQGAKEFGHVOFVGSYKILAYSNDGEHMTYODEKORKDVPFQ 360
 Db 376 sqwlyqvdllvptkvtgltfgaekdfghvqfvgysylaysndgehmtyqdeqkqkdvfg 435
 Qy 361 GNPNDNRKKNVDPPIYARHRIILPMSWYGRITLASFLCTGTEEE 406
 Db 436 gntdndchrknvdppliyarhrlilpwsygritlasellgctee 481
 RESULT 3
 W10364
 ID W10364 standard; Protein: 480 AA.
 AC W10364;
 XX
 DT 03-MAY-1997 (first entry)
 DE Mouse developmentally-regulated endothelial cell locus-1 protein.
 KW Del-1; developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnerable; bone formation; diagnosis; therapy.
 XX
 OS Mus sp.
 FH Key
 FT Peptide
 FT 1..23
 FT /label= Sig_peptide
 FT Protein 24..480
 FT /label= Mat_protein
 FT 26..59
 FT Domain
 FT /label= EGF-1
 FT /note= "epidermal growth factor-like domain 1"
 FT 78..116
 FT /label= EGF-2
 FT /note= "epidermal growth factor-like domain 2"
 FT 123..154
 FT /label= EGF-3
 FT /note= "epidermal growth factor-like domain 3"
 FT 158..314
 FT /label= Discoidin-1
 FT /note= "discoidin 1/factor VIII-like domain 1"
 FT 319..476
 FT /label= Discoidin-2
 FT /note= "discoidin 1/factor VIII-like domain 2"
 XX
 PN W09640769-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 05-JUN-1996; 96WO-US09456.
 XX
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
 XX WPI; 1997-052233/05.
 DR N-PSDB; T47338.

XX
 PT New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis
 XX
 PS Claim 3; Fig 6; 137pp; English.
 XX
 CC Murine Del-1 (W10364) is the polypeptide product of the murine
 CC developmentally-regulated endothelial cell locus-1 (del-1) gene
 CC (747338). It shows 94% amino acid homology to the human Del-1
 CC protein (W10365). Structurally, members of this novel gene family
 CC contain 3 EGF-like domains and 2 discoidin 1/factor VIII-like
 CC domains. Del-1 is expressed in endothelial and certain tumour
 CC cells. Its ability to inhibit vascular formation allows its used
 CC as an anti-angiogenic agent. It can be used as a tumour marker,
 CC to identify Del-1 binding partners, and to modulate endothelial
 CC cell growth and blood vessel formation. Recombinant Del-1 can be
 CC produced in transformed host cells utilizing vectors incorporating
 CC del-1 nucleic acids.
 XX
 SQ Sequence 480 AA;
 Query Match 95.9%; Score 2170.5; DE 18; Length 480;
 Best Local Similarity 95.8%; Pred. No. 1,4e-141;
 Matches 389; Conservative 6; Mismatches 10; Indels 1; Gaps 1;
 Qy 1 GPCTPNFCNNGTCEISEAYRSDTFIGYVCKCPRGFNGIHCOHNIINCEVEPCKNGIGCT 60
 Db 76 gpcitpnchngscetseayrgdtffgyvckcprgfnghcqhnlncceepcngict 135
 Qy 61 DVANVSCSECPGFMGRNOCYKSGPLGEGGISNOGITFASSTHRAFLFGLOKWPYVAR 120
 Db 136 dlvanscsecpgefmrncqkcsghlglegllsngqitassnhralflgkwpypvar 195
 Qy 121 LNKKGILNANTAAENDRMRWITQINLQRRKRVYGVITGAKRISPEYIKFYKIANSNG 180
 Db 196 LnkKjlinawlaeandrnw-rwiqlqrkmrvtgltgskrlyspkfyklynsndg 254
 Qy 181 KTWAMYKVKGTNDMEVFRGNIDNNTPYANSFPPKKAQYVLYQVCRHCTRMELTGC 240
 Db 255 ktwamykvkgtneawfrgnidnntpyansfppkkaqyvllyqvcrrhctrmel1gc 314
 Qy 241 ELSCSEPLGMSKSGHIDYQITASSIFRLNMDMTWEPKRAKLDKQKYNAMTSGHNDQ 300
 Db 315 elsgcseplgmksghidqyqitassifrlmdmtwepkrakldkqgkynawtsgndq 374
 Qy 301 SOWLOYDLVPTKVTGIIITQGAKEFGHVOFVGSYKILAYSNDGEHMTYODEKORKDVPFQ 360
 Db 375 sqwlyqvdllvptkvtgltfgaekdfghvqfvgysylaysndgehmtyqdeqkqkdvfg 434
 Qy 361 GNPNDNRKKNVDPPIYARHRIILPMSWYGRITLASFLCTGTEEE 406
 Db 435 gntdndchrknvdppliyarhrlilpwsygritlasellgctee 480
 RESULT 4
 W94683
 ID W94683 standard; Protein: 480 AA.
 AC W94683;
 XX
 DT 05-MAY-1999 (first entry)
 DE Murine Del-1 protein.
 KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KW discoidin 1; factor VIII-like domain; epidermal growth factor; EGF;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 XX
 OS Mus sp.

```

PN US5877281-A.
XX
XX 02-MAR-1999.
XX
XX 05-JUN-1996; 96US-0659235.
XX
XX 05-JUN-1996; 96US-0659235.
XX
XX 07-JUN-1995; 95US-0480229.
XX
XX (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
XX Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX DR N-PSDB; X18506.
XX
XX Del-1 polypeptide sequences - useful for treatment of cancer,
XX diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX PS Claim 1; Fig 3; 73pp; English.
XX
CC The present sequence is murine developmentally-regulated endothelial
CC cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like
CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
CC have an inhibitory effect on angiogenesis (blood vessel growth), this
CC activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
XX Sequence 480 AA:
SQ
Query Match 95.9%; Score 2170.5; DB 20; Length 480;
Best Local Similarity 95.8%; Pred. No. 1.4e-141;
Matches 389; Conservative 6; Mismatches 10; Indels 1; Gaps 1;
QY 1 GPCPNCHNGTCEISEAYRGDTFTGYGCKPGRGFGNHGCHNINCEVEPCNGGICIT 60
DB 76 gpcipnchngtceiseayrgdtftgygckpgrgfnghcnhineceepcnggic 135
QY 61 DLVANYSCGPCGPEFGMRCQYKCSGPLGIESGIIISNOGITASSSTRALFGLQWYPPYAR 120
DB 136 dlvanyscepcgpefgmrcqykcsgplgiegslisngqiltassnhralfglqkwpypyar 195
QY 121 LNRKGLINAWTAENDRMKRWIOINLQRMRYTGVTGAKRIGSPERIKFYKIAVSDNG 180
DB 196 lnkrglinawtaendrmpwqldnrgmrvtygtgagrkspeyiksykiasvndg 254
QY 181 KTMAMRYVKQINDMVFRCGNIDNTPRYANSFTRPKAQLYALYQVCRHCTLRMELIGC 240
DB 255 ktmamryvkqindmwfrcgnidntrpyansftrpkagvlylpyqtrchctlrmligc 314
QY 241 ELISGCEPLGKMSGHIQDIQTASSIFRTLNMDFTWEPKARLDKQKYNAMTSGHNDQ 300
DB 315 elsgcseplgkmsghiqdyqitassvfrtlnmdftwepkarldkqgkvnawtsghndq 374
QY 301 SOMLOVLLAPTKVTGLITOGAKRDFGHQVFSYKLAISNGCEHWYQOEKQRKQKVP 360
DB 375 sqwlgvllaptkvtglitcgakdfghvqfvsyklaysngcehwmyhqdqekqkdkvlg 434
QY 361 GNFDNDTHRRNVLDPPITVAHRIRLIPWSYGRITLASELLCTEE 406
DB 435 gnfdndthrkuvldppiyarfrilipwsygritlrseellgcaee 480
RESULT 5

```

```

W94697
ID W94697 standard; Protein: 321 AA.
XX
XX AC W94697;
XX
XX DT 05-MAY-1999 (first entry)
XX
XX DE Human milk fat globule protein MFG-E8.
XX
XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
XX diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX
XX OS Homo sapiens.
XX
XX PN US5877281-A.
XX
XX PD 02-MAR-1999.
XX
XX PF 05-JUN-1996; 96US-0659235.
XX
XX PR 05-JUN-1996; 96US-0659235.
XX
XX PR 07-JUN-1995; 95US-0480229.
XX
XX PA (PROG-) PROGENITOR INC.
XX PA (UYVA-) UNIV VANDERBILT.
XX
XX PI Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
XX WPI; 1999-189720/16.
XX
XX DR Del-1 polypeptide sequences - useful for treatment of cancer,
XX PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
XX PS Example: Column 63-64; 73pp; English.
XX
XX CC The present sequence represents human milk fat globule protein MFG-E8,
XX CC which has homology to the developmentally-regulated endothelial cell
XX CC locus 1 (Del-1). The Del-1 protein has epidermal growth factor (EGF) like
XX CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
XX CC have an inhibitory effect on angiogenesis (blood vessel growth), this
XX CC activity may be useful clinically to prevent neovascularisation of
XX CC tissues such as tumour nodules and prevention of metastases. The anti-
XX CC angiogenic activity of Del-1 may be used to treat abnormal conditions
XX CC that result from angiogenesis, including cancer, diabetic retinopathy,
XX CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
XX CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
XX CC stroke, wound healing and peripheral vascular disease. Del-1 is also
XX CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
XX CC integrin, and is an apoptosis inducer.
XX
XX Sequence 321 AA:
SQ
Query Match 75.8%; Score 1717; DB 20; Length 321;
Best Local Similarity 98.4%; Pred. No. 1.2e-110;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 83 CSEPLGIEGIIISNOGITASSSTRALFGLQWYPPYARLNKGLINAWTAENDRMKRWI 142
DB 1 cseplgiegslisngqiltassstralfglqkwpypyarlknkglinawtaendrwmrvi 60
QY 143 QINLQRMRYTGVTGAKRIGSPERIKFYKIAVSDNGKTMAMRYKVTGNDMVFRCGNID 202
DB 61 qnlqgrmrvtygtgagrkspeyikfykiasvndgktmamrykvgtndmwfrcgnid 120
QY 203 NTPRYANSFTRPKAQLYALYQVCRHCTLRMELIGCEPLGKMSGHIQDIQT 262
DB 121 ntrpyansftrpkagvlylpyqtrchctlrmligcseplgkmsghiqdyqit 180
QY 263 ASSIFRTLNMDFTWEPKARLDKQKYNAMTSGHNDQSOVLQDILVPRKYVGIITOGA 322

```

Db 181 asfifltlmamftwepkari|dkgkvnawtsqndsgwlyxllvplkvltgtga 240
 QY 323 KDFGVHVFVSTKLAISNDGEHWIVYODEKOROKVFOGNDTNRKNVDPPIYARHI 382
 Db 241 kdxghvqfvsykhlaynsdghewltvsgdekqrkdkvxgqndthrkknidppiyarhl 300
 QY 383 RILPMGNGRITLASELIGCT 403
 Db 301 rilpmwgyritlaseeligt 321

RESULT 6
 Y94454
 ID Y94454 standard; Protein: 426 AA.
 AC Y94454;
 DT 11-SEP-2000 (first entry)
 DE Mouse lactadherin protein.
 KW Human; lactadherin; MGF-E8; anti-tumour; immune response;
 KW exosome; dendritic cell.
 OS Mus sp.
 FH Key
 FH Peptide 1..22
 FT /label= Secretion_signal
 FT Protein 23..426
 FT /label= Lactadherin
 FT Binding-site 87..89
 FT /label= Integrin_binding_site
 FT Misc-difference 93..111
 FT /note= "encoded by GT"
 PN EP1004664-A1.
 PD 31-MAY-2000.
 PF 24-NOV-1998; 98EP-0402925.
 PR 24-NOV-1998; 98EP-0402925.
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (CUR1-) INST CURIE.
 DR WPI: 2000-352597/31.
 DR N-PSDB; A27141.
 PT Chimeric isolated (human) lactadherin polypeptide that functions as an
 PT adaptor of cross-priming to eliminate pathogenic antigens, e.g. in
 PT tumors -
 PS Disclosure; Page 12; 20pp; English.
 CC Lactadherin protein was found in exosomes produced by dendritic cells.
 CC The protein is involved in the phagocytosis of particulate antigens by
 CC dendritic cells. Exosomes produced by dendritic cells exposed to
 CC tumour antigens induce potent immune responses. Lactadherin or variants
 CC of it may be used in the mediation of an immune response. Variants of
 CC lactadherin may be used for inhibition and/or stimulation of the
 CC cross-priming of antigens and stimulation of the phagocytosis of
 CC antigens by dendritic cells. Compositions derived from lactadherin can
 CC also be used to monitor an immune response, more specifically a CTL
 CC (cytotoxic T-lymphocyte) response and also to produce CTLs specific
 CC for a selected antigen. The present sequence is the mouse
 CC lactadherin protein.
 SQ Sequence 426 AA;

Query Match 46.9%; Score 1061.5; DB 21; Length 426;

Best Local Similarity 49.0%; Pred. No. 1,7e-65;
 Matches 204; Conservative 64; Mismatches 115; Indels 33; Gaps 8;
 QY 3 CTNPCHNGGTC---EISAYRGDFTFYGVCCKPRGFNGIHCOHIN-----CEVEPCK 54
 Db 28 cdselcngtcltgqndily-----clcepfctglvc---neterypcspocy 74
 QY 55 NGGICT-----DLVANSCECEPEFGRNCOYKCGSPGIEGGIISNOQTASTHR 106
 Db 75 ndakclvltlgtqgdiftleyicgpyysgihncelestqlgmeggaladqisasyvm 134
 QY 107 ALFGLQKWPYPYARLNKGLINAWTAENDRWKRMIOQINLQRMKRVGVITOGAKRIGSP 166
 Db 135 gimglqfwpdelarlyrtglvnamhasnyds-kpvlqynllrkmvsgvmqgsaragra 193
 QY 167 EYIKFYKIASNDGKTMAWVKVGTNEDWFRGNDNTNPANSPTPIKAOYRLPYOV 226
 Db 194 eylktfkwaysldgrkfeftgde-sgqdekfignldmslkvmfnplteagylrlpyvs 252
 QY 227 CRRHCTLRMELLCGLSCGSEPIGKMSGHIODYQTASSIFRTLMDMFTWEPKARLDK 286
 Db 253 chrgcltrfclgclngclpigliknltipsqmsasssyktwlltafgyphlgrldn 312
 QY 287 QGKNAMWTSNGHNDQSOQLQVLDLVPKVTGITTQGAKDPGHVQFVGSYKLAISNDGEHWI 346
 Db 313 qgkinawtqagsnakewlydydgtqrgvtgiltgqardtghngyveskvhadsddgywt 372
 QY 347 VYODEKOROKVFOGNDTNRKNVDPPIYARHILPMSWGRITLASELIGC 402
 Db 373 yy--eegsskvtfgnldnshkknfexkrfmayrvlypswhrllrlrltlllgc 426

RESULT 7
 R77252
 ID R77252 standard; Protein: 387 AA.
 AC R77252;
 DT 21-NOV-1995 (first entry)
 DE HMEG 46 kDa antigen.
 KW HMEG; human milk fat globule; antigen; immunogen; vaccine; virucide;
 KW epithelium; tumor; breast cancer; monoclonal antibody; Mab.
 OS Homo sapiens.
 PN W09515171-A.
 PD 08-JUN-1995.
 PF 05-DEC-1994; 94MO-US13967.
 PR 03-DEC-1993; 93US-0162402.
 PA (CANC-) CANCER RES FUND CONTRA COSTA.
 PI Ceriani RL, Larocca DJ, Peterson JA;
 DR WPI: 1995-215151/28.
 DR N-PSDB; Q91198.
 PT 46 kD apparent molecular weight human milk fat globule antigen -
 PT used in assays to determine the presence of a cancerous tumour of
 PT epithelial origin, and in a vaccine against neoplastic tumours
 PS Claim 6; Page 46-47; 68pp; English.
 CC A complete cDNA sequence for the 46 kDa HMEG antigen, a major
 CC component of the apical surface of the normal breast epithelial
 CC cell, was obtained by PCR and RACE methods. cDNA clones can be
 CC used to prepare Mabs for use in immunotherapy, immunohistopathology,
 CC prognosis, diagnosis, imaging and therapy. Recombinant antigen can

Db 76 gpcldpncfhngscatseeysgdftfiygvckcpryflnglhcqhmleceacepcrnrgact 135

Oy 61 dlvanyseegpfgfwmrnokvkgsgplglticgltisnoottasrhrataglokwpyrar 120

Dd 136 dlvanyseecpgeftmrcnqgkcsghlglsagglisnqglstasnmrltgiqkwyryar 195

Oy 121 LKKGLINAMTAENDRMKRWIQTNL 146

Dd 196 lkkkgllnawlaendrnp-wlyqvtv 220

RESULT	11
W94684	
ID	W94684 standard; Protein; 221 AA.

DT	05-MAY-1999 (first entry)
XX	
DE	Truncated murine Del-1 protein.

KM Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KM discoidin I; factor VIII-like domain; epidermal growth factor; EGF
KM diabetic retinopathy; rheumatoid arthritis; endometriosis;
KM angiogenesis.

OS Mus sp.

PN US5877281-A.

PD 02-MAR-1999.

PF 05-JUN-1996; 96US-0659235.

05-JUN-1996: 96US-0659235.

XX 0/ CON ASSOCI 5005-04802225

PA (NYVA-) CIVIL VANDERBILT
FROSTEN LION INC.
(FROG)

PI Hodan B. Quartermoult T. Snodgrass HB. Zundancic TT.

WPY: 1999-189720/16

W-ESUB; ALB07.
XX

Del-1 polypeptide sequences - useful for treatment of cancer

XX
25
CJ 218 31 201:ms 71-73 73-83
FBI/DOJ

The present sequence is truncated murine developmentally-regulated endothelial cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1 proteins have an inhibitory effect on angiogenesis (blood vessel growth). This activity may be useful clinically to prevent neovascularisation of tissues such as tumour nodules and prevention of metastases. The anti-angiogenic activity of Del-1 may be used to treat abnormal conditions that result from angiogenesis, including cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be used to treat cardiac ischaemia, thrombotic stroke, wound healing and peripheral vascular disease. Del-1 is also useful for promoting bone formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis inducer.

Sequence 221 AA;

Query Match	34.7%;	Score 784.5;	DB 20;	Length 221;
Best Local Similarity	92.58;	Pred. No. 8.7e-47;		
Matches 135; Conservative	3;	Mismatches 7;	Indels 1;	Gaps 1;

QY 1 GPCTPNCNNGGTCEISEAYRGDTFIGYVCKPKRPFNGIHCQHINIECEEVEPCKNIGICT 600
||| |||||||||||||||||||||||||||||||||||||||| |::|||

Db 76 gpeipnphmgagccseayrgdtfigyvkpcprgfnghqgmhinecaeaperngdtc 135a

Qy 61 dlvamvseecpgepkmrnoqkccspglctgctgllnmoqtllsssthralfgldokmtpyrr 120a

Db 136 dlvayaseecpgegmrmcnygcspghglsagglllnqglatssmhralfgldkwyryaa 195a

Qy 121 lnrkgllnwmvtaendrmrkwrtoinl 146a

Db 196 lnhkylgnawtaeandrpv-wlqytlv 220a

RESULT	12
B03533	
ID	B03533 standard; Protein: 2183 AA.

DT	09-OCT-2000	(first entry)
XX		
DE	Murine factor V SEQ ID NO:5.	

KM Murine; factor V; FV; activated protein C; APC; anticoagulant;
KM activated protein C resistant factor V; thrombosis; screening
KM thrombophilia.
XX

XX of

[illegible]

XX 2
C
A
A
C
C

	X	X	F
	X	X	F
		C	
		C	
		Z	
		C	
		A	
		H	
		V	
		C	
		V	
		C	
		E	
		C	
		A	
		C	
		H	
		T	

00 NOV - 1530, 3005-0/40111
FX
XX

PA (UNML) UNIV MICHIGAN.
YY

PL Ginsburg D, Cul J
VY

DR WPL; 2000-410682/35
DR N-BEND; 260446

XX
XX
New Technology

PT and factor V null transgenic mice useful for screening anticoagulants

gene therapy protocols .

PS Example 1; Fig 2; 76pp; English.

The present invention describes transgenic mice (I) and (II) containing modifications in the factor V gene, where (I) expresses an activated protein C (APC) resistant factor V, and (II) lacks the ability to express wild-type factor V. The transgenic animals (I) and (II) are useful for screening compounds with anticoagulant activity. Methods from the present invention, and the transgenic animals, are also useful in providing models for human thrombophilia. These models are useful in providing insight into the basic regulatory mechanisms of blood coagulation and pathogenesis of human thrombosis. In addition, factor V null transgenic mice, especially pregnant females may be used as a model system to test *in utero* gene replacement therapy protocols. The present sequence represents murine factor V, which is used in an example from the present invention.

SQ Sequence 2183 AA;

Query Match	30.1%	Score 682.5;	DB 21;	Length 2183;
Best Local Similarity	43.2%;	Pred. No. 1.1e-38;		
Matches 140; Conservative	51;	Mismatches 122;	Indels 11;	Gaps 3;

Qy 82 KCSGPLEGGIISNQLTASSTHRAFLGLQKWPYPYARLNKKGLINAMT--AAENDRW 138
: | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 1865 eckmmpmqstlyvtsdsqkrlasey-----ltwepriarlmaagsynawsiektaldfdpl 1918

QY 139 KRMIQINLQKRMKRVGTITOGAKRIGSPYIKFKIAYSNQKTMWYKVKGTEDWVFR 198
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1919 kpwlyqvmqevvvtqitqgahylksctteitqvayssqdtmqitryqskysvmyft 1978
QY 199 GNIDNNTPVYNSFTPIKAQYVRLYPQVCRHRTLMELLGCELSGCEPRLGMSGHIOD 258
| | : | : | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1979 gnsdgsitkenrlidpplivaritrlhpcksynprrlrllelqgeevngscstplqledgrlqd 2038
QY 259 YQITASSIFPTLMDMFTEPRKARLDKQGVNMTSGHNSQSMLOVDLIVPTKVTGII 318
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 2039 kqitassfkkswwdy--wepslarluaqrvnawqakannkqylqvdllkikkvcaiv 2096
QY 319 TQGAQDFGHQVFGSYKLAIVSNDGEHMTYODEKQKRDVFOGNEFDNDTHRKNVIDPPY 378
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 2097 tlgckslssemyvkyslyqsdqgvawkpyrqkssmvdklilegnantkghnknfnppl 2156
QY 379 ARHRIILPMWGRITLASHLGC 402
| | | | | : | | | | : | | | | : | | | | : | | | | :
Db 2157 srflrlipkctmgslarllelqgc 2180

RESULT 13
W04254
ID W04254 standard; Protein: 2224 AA.
XX
AC W04254;
XX
DT 14-DEC-1996 (first entry)
XX
DE Human Factor V.
XX
KW Factor V; activated Protein C resistance; APC; genetic screening;
XX
KW allele; point mutation; diagnosis.
XX
OS Homo sapiens.
XX
FH Key
FT 1..28 Location/Qualifiers
FT Peptide
FT /label= Sig_peptide
FT 29..736
FT Region
FT /label= Heavy-chain_region
FT 737..1573
FT Region
FT /label= Connecting_region
FT 1573..2224
FT Region
FT /label= Light_chain_region
FT 51
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 55
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 239
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 297
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 382
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 460
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 468
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 535
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 667
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 737..738
FT Modified-site
FT /note= "thrombin cleavage site"
FT 742
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 753
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 760
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 776
FT Modified-site
FT /note= "potential N-glycosylation site"

FT Modified-site
FT /note= "potential N-glycosylation site"
FT 782
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 821
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 938
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 976
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1046..1047
FT Modified-site
FT /note= "thrombin cleavage site"
FT 1075
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1083
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1103
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1106
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1203
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1221
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1257
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1266
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1293
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1311
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1338
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1347
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1374
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1383
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1479
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1489
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1539
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 1573..1574
FT Modified-site
FT /note= "thrombin cleavage site"
FT 1703
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 2010
FT Modified-site
FT /note= "potential N-glycosylation site"
FT 2209
FT Modified-site
FT /note= "potential N-glycosylation site"

W09630546-A1.
03-OCT-1996.
XX
22-MAR-1996; 96MO-US03881.
XX
24-MAR-1995; 95US-0410488.
XX
(SCRI) SCRIPPS RES INST.
XX
Gandrille S, Greengard J, Griffin JH.
XX
WPT. 1996-455389/45.
XX
DR N-PSDB; T33942.
XX
Detection of Factor V gene mutation - by PCR amplification to
XX identify exon 10 guanine 205 or 1691 to adenine substitution, which
XX results in activated Protein C resistance

PA (UYEM-) UNIV EMORY.

PI Lollar JS, Runge MS;

DR WPI; 1998-271107/24.
N-DCDD; W3E913

DR N-PSDB; V25812.

PT Hybrid of human and animal factor VII - containing porcine and
PT murine amino acid sequences is useful in the treatment of
PT haemophilia

PS Claim 9; Column 73-84; 48pp; English.

The present sequence represents murine factor VIII used in the present invention. The present invention describes a new procoagulant hybrid factor VIII (I) comprising human factor VIII and has amino acid sequences substituted from the group of A2 domain fragments consisting of amino acids 373-540, 373-508, 445-508, 404-508, 489-508 and 484-489 from the human 2332 amino acid A2 domain sequence (II) as given in the specification (see w53483). The substitution is from corresponding non-human mammalian factor VIII sequences. Also described is a method for treating factor VIII deficiency comprising administering a therapeutically effective dose of (I) in a pharmaceutical carrier. (I), prepared from reconstitution of purified molecules or recombinant techniques, is useful in the treatment of haemophiliacs who have factor VIII deficiencies and whose blood is not normally able to clot after internal or external bleeding. (I) compared to native human factor VIII is more stable at physiological conditions and has a higher specific clotting activity.

Sequence 2319 AA;

Query Match	29.6%;	Score 670;	DB 19;	Length 2319;
Best Local Similarity	46.3%;	Pred. No. 8.1e-38;		
Matches 151;	Conservative 40;	Mismatches 119;	Indels 16;	Gaps 5,

[illegible]

Search completed: May 23, 2001, 06:15:17
Job time: 231 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:18:14 ; Search time 179.67 Seconds
(without alignments)
363.483 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513
Perfect score: 2264
Sequence: 1 GPCTPNCHNGTGCTEISEAY.....WSWYGRITLSELLGCTEEE 406

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues
Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/PCRTUS.COMB.pep: *
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep: *
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep: *
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep: *
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep: *
6: /cgn2_6/ptodata/2/paa/US083.COMB.pep: *
7: /cgn2_6/ptodata/2/paa/US084.COMB.pep: *
8: /cgn2_6/ptodata/2/paa/US085.COMB.pep: *
9: /cgn2_6/ptodata/2/paa/US086.COMB.pep: *
10: /cgn2_6/ptodata/2/paa/US087.COMB.pep: *
11: /cgn2_6/ptodata/2/paa/US088.COMB.pep: *
12: /cgn2_6/ptodata/2/paa/US089.COMB.pep: *
13: /cgn2_6/ptodata/2/paa/US090.COMB.pep: *
14: /cgn2_6/ptodata/2/paa/US091.COMB.pep: *
15: /cgn2_6/ptodata/2/paa/US092.COMB.pep: *
16: /cgn2_6/ptodata/2/paa/US093.COMB.pep: *
17: /cgn2_6/ptodata/2/paa/US094.COMB.pep: *
18: /cgn2_6/ptodata/2/paa/US095.COMB.pep: *
19: /cgn2_6/ptodata/2/paa/US096.COMB.pep: *
20: /cgn2_6/ptodata/2/paa/US097.COMB.pep: *
21: /cgn2_6/ptodata/2/paa/US098.COMB.pep: *
22: /cgn2_6/ptodata/2/paa/US099.COMB.pep: *
23: /cgn2_6/ptodata/2/paa/US100.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2264	100.0	513	8	US-08-480-229B-14
2	2264	100.0	513	16	US-08-659-235B-14
3	2264	100.0	513	16	US-09-237-981-14
4	2235	98.7	481	10	US-08-659-235B-29
5	2170.5	95.9	480	8	US-08-480-229B-10
6	2170.5	95.9	480	10	US-08-659-235B-10
7	2170.5	95.9	480	16	US-09-237-981-10
8	1717	75.8	321	8	US-08-480-229B-21
9	1717	75.8	321	10	US-08-659-235B-21
10	1717	75.8	321	16	US-09-237-981-21

11	1615	71.3	449	23	US-60-207-315-404	Sequence 404, App
12	1406.5	62.1	362	23	US-60-207-315-523	Sequence 523, App
13	1405.5	62.1	311	23	US-60-230-435-1519	Sequence 1519, App
14	1075	47.5	463	19	US-09-582-340-4	Sequence 4, Appl
15	1020.5	45.1	387	19	US-09-582-340-2	Sequence 2, Appl
16	948	41.9	320	8	US-08-480-229B-20	Sequence 20, Appl
17	948	41.9	320	10	US-08-659-235B-20	Sequence 20, Appl
18	948	41.9	320	16	US-09-237-981-20	Sequence 20, Appl
19	816.5	36.1	203	23	US-60-233-644-77	Sequence 77, Appl
20	813.5	35.9	185	23	US-60-230-435-1806	Sequence 1806, App
21	784.5	34.7	221	16	US-09-237-981-29	Sequence 29, Appl
22	675.5	29.8	2224	8	PCT-US96-03881A-14	Sequence 14, Appl
23	675.5	29.8	2224	14	US-08-410-488-14	Sequence 14, Appl
24	670.5	29.6	2224	23	US-60-208-020-170	Sequence 170, App
25	670.5	29.6	2228	23	US-60-209-043-221	Sequence 221, App
26	670.5	29.6	2228	23	US-60-213-172-167	Sequence 167, App
27	670.5	29.6	2228	23	US-60-213-173-167	Sequence 6, Appl
28	670.5	29.6	2319	1	PCT-US01-13541-6	Sequence 28, Appl
29	670	29.6	2319	1	PCT-US01-05076A-28	Sequence 6, Appl
30	670	29.6	2319	1	PCT-US97-11155-6	Sequence 6, Appl
31	670	29.6	2319	1	PCT-US99-05193-6	Sequence 6, Appl
32	670	29.6	2319	8	US-08-475-201-6	Sequence 6, Appl
33	670	29.6	2319	17	US-09-315-179-6	Sequence 6, Appl
34	670	29.6	2319	19	US-09-523-656-28	Sequence 28, Appl
35	670	29.6	2304	14	US-09-035-141-4	Sequence 4, Appl
36	655.5	29.0	2304	17	US-09-324-867-4	Sequence 1, Appl
37	655.5	29.0	2304	1	PCT-US99-29169-1	Sequence 1, Appl
38	649	28.7	1438	16	US-09-209-916-1	Sequence 16, Appl
39	649	28.7	1438	12	US-08-869-309-47	Sequence 47, Appl
40	649	28.7	1457	14	US-09-001-039A-47	Sequence 47, Appl
41	649	28.7	1457	14	US-09-001-039B-47	Sequence 47, Appl
42	649	28.7	1457	15	US-09-190-941-47	Sequence 47, Appl
43	649	28.7	1471	10	US-08-683-838A-3	Sequence 3, Appl
44	649	28.7	1661	16	US-09-243-532-2	Sequence 2, Appl
45	649	28.7				

ALIGNMENTS

RESULT 1
US-08-480-229B-14
; Sequence 14, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quartermoss, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229B-14

Query Match 100.0%; Score 2264; DB 8; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.7e-201;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCPNCHNGTCEISEAYRGDTFTGYVCKPRGFNGHCOHINNECEVEPCCKNGICT 60
DB 108 GPCPNCHNGTCEISEAYRGDTFTGYVCKPRGFNGHCOHINNECEVEPCCKNGICT 167
QY 61 DLVANYSCCEPPEFMGRNCOYKCSGPLGIEGIIISNOQITASTHRALFGLQKWPYYAR 120
DB 168 DLVANYSCCEPPEFMGRNCOYKCSGPLGIEGIIISNOQITASTHRALFGLQKWPYYAR 227
QY 121 LNKGLINAMTAENDRMKRWIOINLQRMKRVYITOGAKRIGSEYIKFYKIAYSDNG 180
DB 228 LNKGLINAMTAENDRMKRWIOINLQRMKRVYITOGAKRIGSEYIKFYKIAYSDNG 287
QY 181 KTWAMVYKGTNEDMVRGNDNNTPYANSFTPIKAQYVRLPYOVCRHCTLRMELGCG 240
DB 288 KTWAMVYKGTNEDMVRGNDNNTPYANSFTPIKAQYVRLPYOVCRHCTLRMELGCG 347
QY 241 ELSCSEPLGMSGHIDYQITASSIFRTLNMDFTEPRKARLDKQKVNAWMTSGHNDQ 300
DB 348 ELSCSEPLGMSGHIDYQITASSIFRTLNMDFTEPRKARLDKQKVNAWMTSGHNDQ 407
QY 301 SOMLOYDLVPTKYVGIITOGAKDFGHVQFVGSYKLAISNDEHNTVYODEKQKDKVFO 360
DB 408 SOMLOYDLVPTKYVGIITOGAKDFGHVQFVGSYKLAISNDEHNTVYODEKQKDKVFO 467
QY 361 GNFNDTHRKNVIDPPIYARHRIPLPMSWYGRITLASELLGCTEE 406
DB 468 GNFNDTHRKNVIDPPIYARHRIPLPMSWYGRITLASELLGCTEE 513

RESULT 2
US-08-659-235B-14
Sequence 14, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quatermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-14

Query Match 100.0%; Score 2264; DB 10; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.7e-201;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GPCPNCHNGTCEISEAYRGDTFTGYVCKPRGFNGHCOHINNECEVEPCCKNGICT 60
DB 108 GPCPNCHNGTCEISEAYRGDTFTGYVCKPRGFNGHCOHINNECEVEPCCKNGICT 167
QY 61 DLVANYSCCEPPEFMGRNCOYKCSGPLGIEGIIISNOQITASTHRALFGLQKWPYYAR 120
DB 168 DLVANYSCCEPPEFMGRNCOYKCSGPLGIEGIIISNOQITASTHRALFGLQKWPYYAR 227
QY 121 LNKGLINAMTAENDRMKRWIOINLQRMKRVYITOGAKRIGSEYIKFYKIAYSDNG 180
DB 228 LNKGLINAMTAENDRMKRWIOINLQRMKRVYITOGAKRIGSEYIKFYKIAYSDNG 287
QY 181 KTWAMVYKGTNEDMVRGNDNNTPYANSFTPIKAQYVRLPYOVCRHCTLRMELGCG 240
DB 288 KTWAMVYKGTNEDMVRGNDNNTPYANSFTPIKAQYVRLPYOVCRHCTLRMELGCG 347
QY 241 ELSCSEPLGMSGHIDYQITASSIFRTLNMDFTEPRKARLDKQKVNAWMTSGHNDQ 300
DB 348 ELSCSEPLGMSGHIDYQITASSIFRTLNMDFTEPRKARLDKQKVNAWMTSGHNDQ 407
QY 301 SOMLOYDLVPTKYVGIITOGAKDFGHVQFVGSYKLAISNDEHNTVYODEKQKDKVFO 360
DB 408 SOMLOYDLVPTKYVGIITOGAKDFGHVQFVGSYKLAISNDEHNTVYODEKQKDKVFO 467
QY 361 GNFNDTHRKNVIDPPIYARHRIPLPMSWYGRITLASELLGCTEE 406
DB 468 GNFNDTHRKNVIDPPIYARHRIPLPMSWYGRITLASELLGCTEE 513

RESULT 3
US-09-237-981-14
Sequence 14, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quatermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

Wed May 23 06:29:17 2001

us-09-237-981-14_copy_108_513.rapm

Page 3

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SRO ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-14
```

Query Match 100.0%; Score 2264; DB 16; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.7e-201;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GPCPNPCHNGGCEISEAYRGDTFPGYCKCPRGFGNGHCHQHNINCEVEPCCKNGIGCT 60
   |||||||
DB 108 GPCPNPCHNGGCEISEAYRGDTFPGYCKCPRGFGNGHCHQHNINCEVEPCCKNGIGCT 167

QY 61 DLVANYSCGCPGEMRNCOYKCSGPLGIEGIIISNOOITASSTHRALFGLQKWPYYAR 120
   |||||||
DB 168 DLVANYSCGCPGEMRNCOYKCSGPLGIEGIIISNOOITASSTHRALFGLQKWPYYAR 227

QY 121 LNKKGILNAAWTAENDRMKRWIOLNLRKMRVTGVTGAKRRIGSPYIKFYKIAVSDG 180
   |||||||
DB 228 LNKKGILNAAWTAENDRMKRWIOLNLRKMRVTGVTGAKRRIGSPYIKFYKIAVSDG 287

QY 181 KTMAMYKVGKGNEDMVRFGNIDNNTPYANSFTPIKAQYRLYPQVCRHCTLRMELLCG 240
   |||||||
DB 288 KTMAMYKVGKGNEDMVRFGNIDNNTPYANSFTPIKAQYRLYPQVCRHCTLRMELLCG 347

QY 241 ELSCGSEPLGKSGHIDYQITASSIFRTLNDMFTWEPKRALDQKGVNMTSGHNDQ 300
   |||||||
DB 348 ELSCGSEPLGKSGHIDYQITASSIFRTLNDMFTWEPKRALDQKGVNMTSGHNDQ 407

QY 301 SQWLQVLDLVPKTVGTGIIITOGAKDFGHVQFVGSYKLAISNDGEHMTVYODEKORDKVFQ 360
   |||||||
DB 408 SQWLQVLDLVPKTVGTGIIITOGAKDFGHVQFVGSYKLAISNDGEHMTVYODEKORDKVFQ 467

QY 361 GNFNDPDRKKNVDPPIYARHRIILPMSWYGRITLASLGLCTEE 406
   |||||||
DB 468 GNFNDPDRKKNVDPPIYARHRIILPMSWYGRITLASLGLCTEE 513
```

RESULT 4
US-08-659-235B-29
Sequence 29, Application US/08659235B

```

GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
```

```

STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SRO ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-29
```

Query Match 98.7%; Score 2235; DB 10; Length 481;
Best Local Similarity 99.0%; Pred. No. 7.5e-199;
Matches 402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 GPCPNPCHNGGCEISEAYRGDTFPGYCKCPRGFGNGHCHQHNINCEVEPCCKNGIGCT 60
   |||||||
DB 76 GPCPNPCHNGGCEISEAYRGDTFPGYCKCPRGFGNGHCHQHNINCEVEPCCKNGIGCT 135

QY 61 DLVANYSCGCPGEMRNCOYKCSGPLGIEGIIISNOOITASSTHRALFGLQKWPYYAR 120
   |||||||
DB 136 DLVANYSCGCPGEMRNCOYKCSGPLGIEGIIISNOOITASSTHRALFGLQKWPYYAR 195

QY 121 LNKKGILNAAWTAENDRMKRWIOLNLRKMRVTGVTGAKRRIGSPYIKFYKIAVSDG 180
   |||||||
DB 196 LNKKGILNAAWTAENDRMKRWIOLNLRKMRVTGVTGAKRRIGSPYIKFYKIAVSDG 255

QY 181 KTMAMYKVGKGNEDMVRFGNIDNNTPYANSFTPIKAQYRLYPQVCRHCTLRMELLCG 240
   |||||||
DB 256 KTMAMYKVGKGNEDMVRFGNIDNNTPYANSFTPIKAQYRLYPQVCRHCTLRMELLCG 315

QY 241 ELSCGSEPLGKSGHIDYQITASSIFRTLNDMFTWEPKRALDQKGVNMTSGHNDQ 300
   |||||||
DB 316 ELSCGSEPLGKSGHIDYQITASSIFRTLNDMFTWEPKRALDQKGVNMTSGHNDQ 375

QY 301 SQWLQVLDLVPKTVGTGIIITOGAKDFGHVQFVGSYKLAISNDGEHMTVYODEKORDKVFQ 360
   |||||||
DB 376 SQWLQVLDLVPKTVGTGIIITOGAKDFGHVQFVGSYKLAISNDGEHMTVYODEKORDKVFQ 435

QY 361 GNFNDPDRKKNVDPPIYARHRIILPMSWYGRITLASLGLCTEE 406
   |||||||
DB 436 GNFNDPDRKKNVDPPIYARHRIILPMSWYGRITLASLGLCTEE 481
```

RESULT 5
US-08-480-229B-10
Sequence 10, Application US/08480229B

```

GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States
;; ZIP: 10036-2711
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/480,229B
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-026
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 Pennie
;;
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 480 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-480-229B-10

Query Match 95.9% Score 2170.5; DB: 8; Length 480;
Best Local Similarity 95.8%; Pred. No. 7.3e-193;
Matches 389; Conservative 6; Mismatches 10; Indels 1; Gaps 1;
;;
QY 1 GPCIPNCHNGTCEISEAYRGDTFGYVCKCPGFNGIHCOHININECEVEPCCKNGICT 60
DB 76 GPCIPNCHNGTCEISEAYRGDTFGYVCKCPGFNGIHCOHININECEVEPCCKNGICT 135
QY 61 DLVANSCEGCEGFMRNCQYKCSGPIGIEGGIISNOQITASSNHRALFGLQKWPYYAR 120
DB 136 DLVANSCEGCEGFMRNCQYKCSGPIGIEGGIISNOQITASSNHRALFGLQKWPYYAR 195
QY 121 LNKKGILNMTAENDRMKRWIOLNORKMRYTGVITOGAKRIGSPETIKFYKIAYSNDG 180
DB 196 LNKKGILNMTAENDRMKRWIOLNORKMRYTGVITOGAKRIGSPETIKFYKIAYSNDG 254
QY 181 KTWAMKYKGTNEWVRGNDNNTPYANSFTPIKAQVRLYPQVCRHCTLMELLGC 240
DB 255 KTWAMKYKGTNEWVRGNDNNTPYANSFTPIKAQVRLYPQVCRHCTLMELLGC 314
QY 241 ELSCGSEPLGKMSGHIDYQITASSVFTLNMDFTWERRKRLDQKGNAMTSGHNDQ 300
DB 315 ELSCGSEPLGKMSGHIDYQITASSVFTLNMDFTWERRKRLDQKGNAMTSGHNDQ 374
QY 301 SOMLOVDLVPRTKVTGIIITOGAKDFGHVQVGSYKLAYSNDGEHMTVYQDEKORDKVFQ 360
DB 375 SOMLOVDLVPRTKVTGIIITOGAKDFGHVQVGSYKLAYSNDGEHMTVYQDEKORDKVFQ 434
QY 361 GNFNDNTHRKNNVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 406
DB 435 GNFNDNTHRKNNVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 480
;;
RESULT 6
US-08-659-235B-10
; Sequence 10, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph

;; APPLICANT: Zupancic, Thomas J.
;; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States
;; ZIP: 10036-2711
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/659,235B
;; FILING DATE: 05-JUN-1996
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Poissant, Brian M.
;; REGISTRATION NUMBER: 28,462
;; REFERENCE/DOCKET NUMBER: 8907-034
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 Pennie
;;
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 480 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-659-235B-10

Query Match 95.9% Score 2170.5; DB: 10; Length 480;
Best Local Similarity 95.8%; Pred. No. 7.3e-193;
Matches 389; Conservative 6; Mismatches 10; Indels 1; Gaps 1;
;;
QY 1 GPCIPNCHNGTCEISEAYRGDTFGYVCKCPGFNGIHCOHININECEVEPCCKNGICT 60
DB 76 GPCIPNCHNGTCEISEAYRGDTFGYVCKCPGFNGIHCOHININECEVEPCCKNGICT 135
QY 61 DLVANSCEGCEGFMRNCQYKCSGPIGIEGGIISNOQITASSNHRALFGLQKWPYYAR 120
DB 136 DLVANSCEGCEGFMRNCQYKCSGPIGIEGGIISNOQITASSNHRALFGLQKWPYYAR 195
QY 121 LNKKGILNMTAENDRMKRWIOLNORKMRYTGVITOGAKRIGSPETIKFYKIAYSNDG 180
DB 196 LNKKGILNMTAENDRMKRWIOLNORKMRYTGVITOGAKRIGSPETIKFYKIAYSNDG 254
QY 181 KTWAMKYKGTNEWVRGNDNNTPYANSFTPIKAQVRLYPQVCRHCTLMELLGC 240
DB 255 KTWAMKYKGTNEWVRGNDNNTPYANSFTPIKAQVRLYPQVCRHCTLMELLGC 314
QY 241 ELSCGSEPLGKMSGHIDYQITASSVFTLNMDFTWERRKRLDQKGNAMTSGHNDQ 300
DB 315 ELSCGSEPLGKMSGHIDYQITASSVFTLNMDFTWERRKRLDQKGNAMTSGHNDQ 374
QY 301 SOMLOVDLVPRTKVTGIIITOGAKDFGHVQVGSYKLAYSNDGEHMTVYQDEKORDKVFQ 360
DB 375 SOMLOVDLVPRTKVTGIIITOGAKDFGHVQVGSYKLAYSNDGEHMTVYQDEKORDKVFQ 434
QY 361 GNFNDNTHRKNNVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 406
DB 435 GNFNDNTHRKNNVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 480
;;
RESULT 7
US-09-237-981-10
; Sequence 10, Application US/09237981

GENERAL INFORMATION:
APPLICANT: Quertemus, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-10

Query Match 95.9%; Score 2170.5; DB 16; Length 480;
Best Local Similarity 95.8%; Pred. No. 7,3e-193;
Matches 389; Conservative 6; Mismatches 10; Indels 1; Gaps 1;
QY 1 GPCTPNCHNGTCEISBAYRGTFTGYCKCPRGFNGIHCOHINCEVEPCKNGICT 60
DB 76 GPCIPNCHNGTCEISBAYRGTFTGYCKCPRGFNGIHCOHINCEVEPCKNGICT 135
QY 61 DLVANYSCDEPFEMGRNCOYKCSGPIEGGIISNOQITASSSTRALFGLQWYVYAR 120
DB 136 DLVANYSCDEPFEMGRNCOYKCSGPIEGGIISNOQITASSSTRALFGLQWYVYAR 195
QY 121 LNKGLINAMTAENDRMKRMIOINLQRMKRYGVITOGAKRISPEYIFYTIAVSNIG 180
DB 196 LNKGLINAMTAENDRMKRMIOINLQRMKRYGVITOGAKRISPEYIFYTIAVSNIG 254
QY 181 KTMAYKYKGTNEDMVRFGNDNTPYANSFTPIKAQYVRLPOVCRHCTLRMELLCG 240
DB 255 KTMAYKYKGTNEDMVRFGNDNTPYANSFTPIKAQYVRLPOVCRHCTLRMELLCG 314
QY 241 ELSCGSEPLGKMSGHIDQYOTASSIFRTLNDMTWEPKARLDKOGKVNAMTSGHND 300
DB 315 ELSCGSEPLGKMSGHIDQYOTASSIFRTLNDMTWEPKARLDKOGKVNAMTSGHND 374
QY 301 SQMLQVDLVPTKYVGIITOGAKDFGVQVSGYKLAASNDGEHMTYODEKORKDKVFO 360
DB 375 SQMLQVDLVPTKYVGIITOGAKDFGVQVSGYKLAASNDGEHMTYODEKORKDKVFO 434
QY 361 GNFNDNTHRKKNVIDPPIYARIRILPMSWGRITLASELCTGTEE 406

DB 435 GNFNDNTHRKKNVIDPPIYARIRILPMSWGRITLASELCTGTEE 480
RESULT 8
US-08-480-229B-21
Sequence 21, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertemus, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-229B-21

Query Match 75.8%; Score 1717; DB 8; Length 321;
Best Local Similarity 98.4%; Pred. No. 5.6e-151;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 83 CSGPIEGGIISNOQITASSSTRALFGLQWYVYARLKKKLLINAMTAENDRMKMI 142
DB 1 CSGPIEGGIISNOQITASSSTRALFGLQWYVYARLKKKLLINAMTAENDRMKMI 60
QY 143 QINQRMKRYGVITOGAKRISPEYIFYTIAVSNIGKTMAYKYKGTNEDMVRFGND 202
DB 61 QINQRMKRYGVITOGAKRISPEYIFYTIAVSNIGKTMAYKYKGTNEDMVRFGND 120
QY 203 NNTPYANSFTPIKAQYVRLPOVCRHCTLRMELLCGSEPLGKMSGHIDQYIT 262
DB 121 NNTPYANSFTPIKAQYVRLPOVCRHCTLRMELLCGSEPLGKMSGHIDQYIT 180
QY 263 ASSIFRTLNDMTWEPKARLDKOGKVNAMTSGHNDQSMLOVDLVPTKYVGIITOGA 322
DB 181 ASSIFRTLNDMTWEPKARLDKOGKVNAMTSGHNDQSMLOVDLVPTKYVGIITOGA 240
QY 323 KDFGVQVSGYKLAASNDGEHMTYODEKORKDKVFOGNDTHRKKNVIDPPIYARI 382
DB 241 KDFGVQVSGYKLAASNDGEHMTYODEKORKDKVFOGNDTHRKKNVIDPPIYARI 300

QY 383 RILPMSWGRTTLASELCT 403
|||||
Db 301 RILPMSWGRTTLASELCT 321

RESULT 9

US-08-659-235B-21
; Sequence 21, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penile & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Polissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Penile
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-659-235B-21

Query Match 75.8%; Score 1717; DB 10; Length 321;
Best Local Similarity 98.4%; Pred. No. 5, 6e-151;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 83 CSGPLGIEGGIISNOQITASTSRHALFGLQKWPYARLNKGLINAWTAENDRNKRWI 142
Db 1 CSGPLGIEGGIISNOQITASTSRHALFGLQKWPYARLNKGLINAWTAENDRNKRWI 60
QY 143 QINLQKRMVTVGTGACARRISPEYIKFYKIAVSDNGKTWAMYKVGTVEDNVPFGNID 202
Db 61 QINLQKRMVTVGTGACARRISPEYIKFYKIAVSDNGKTWAMYKVGTVEDNVPFGNID 120
QY 203 NNTPYANSFTPIKAQVRLYPQVCRRHCTLMELLGCELSGCESEPLGKMSGHIDYQIT 262
Db 121 NNTPYANSFTPIKAQVRLYPQVCRRHCTLMELLGCELSGCESEPLGKMSGHIDYQIT 180
QY 263 ASSTFRILNMDFTWEPKARLDKQKVNAMTSGHNDOSQMLQVLLVPYKVTGIIITGGA 322
Db 181 ASSTFRILNMDFTWEPKARLDKQKVNAMTSGHNDOSQMLQVLLVPYKVTGIIITGGA 240
QY 323 KDFGHVDFVGSYKLAISNDGEHMTYXQDEKQKDKVAGNFDNDTHRKAVIDPPIYARHI 382
Db 241 KDFGHVDFVGSYKLAISNDGEHMTYXQDEKQKDKVAGNFDNDTHRKAVIDPPIYARHI 300

QY 383 RILPMSWGRTTLASELCT 403
|||||
Db 301 RILPMSWGRTTLASELCT 321

RESULT 10

US-09-237-981-21
; Sequence 21, Application US/09237981
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penile & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/237,981
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235
; FILING DATE: 05-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Polissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Penile
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-237-981-21

Query Match 75.8%; Score 1717; DB 16; Length 321;
Best Local Similarity 98.4%; Pred. No. 5, 6e-151;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 83 CSGPLGIEGGIISNOQITASTSRHALFGLQKWPYARLNKGLINAWTAENDRNKRWI 142
Db 1 CSGPLGIEGGIISNOQITASTSRHALFGLQKWPYARLNKGLINAWTAENDRNKRWI 60
QY 143 QINLQKRMVTVGTGACARRISPEYIKFYKIAVSDNGKTWAMYKVGTVEDNVPFGNID 202
Db 61 QINLQKRMVTVGTGACARRISPEYIKFYKIAVSDNGKTWAMYKVGTVEDNVPFGNID 120
QY 203 NNTPYANSFTPIKAQVRLYPQVCRRHCTLMELLGCELSGCESEPLGKMSGHIDYQIT 262
Db 121 NNTPYANSFTPIKAQVRLYPQVCRRHCTLMELLGCELSGCESEPLGKMSGHIDYQIT 180
QY 263 ASSTFRILNMDFTWEPKARLDKQKVNAMTSGHNDOSQMLQVLLVPYKVTGIIITGGA 322
Db 181 ASSTFRILNMDFTWEPKARLDKQKVNAMTSGHNDOSQMLQVLLVPYKVTGIIITGGA 240

QY 323 KDGHVQFVGSYKLYASNDGEHMYVYODEKORCKVPOGPNFNTDTHKKNVIDPEIYARHT 382
DB 241 KDXGHVQFVGSYKLYASNDGEHMYVYODEKORCKVXGPNFNDTHKKNVIDPEIYARHT 300
QY 383 RILPMSWGRITLASELLGCT 403
DB 301 RILPMSWGRITLASELLGCT 321

RESULT 11

US-60-207-315-404
; Sequence 404, Application us/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NOCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 404
; LENGTH: 449
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(449)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-207-315-404

Query Match 71.3%; Score 1615; DB 23; Length 449;
Best Local Similarity 70.6%; Pred. No. 2,6e-141;
Matches 305; Conservative 2; Mismatches 7; Indels 118; Gaps 3;

QY 2 PCPNPCHNGCTGEISEAYRGDTFYGVCRCRFGNGIHCOH----- 43
DB 2 PCPNPCHNGCTGEISEAYRGDTFYGVCRCRFGNGIHCOHRLSHSCPALADPLSVPG 61
QY 44 ----- 43
DB 62 KPLPLSLIDNAPRRSSKNAPIRISTSSGRKVPPLSTLREPSDWEHVGPKAAMNTGLAMPVL 121
QY 44 ----- 86
DB 122 RGGILGOLLRPECGAPNINECEVEPCCKNGGICTDLVANYSCGCEPFGRNCOYKCSGP 181
QY 87 LGIEGGIISNOQITASTHRALEGLQKWYPYARLNKKGLINAMTAENDRWK----- 139
DB 182 LGIEGGIISNOQITASTHRALEGLQKWYPYARLNKKGLINAMTAENDRWK----- 139
QY 140 -----RWIQLNLRKRMRYTGVIITOGAKRISPEYIKFYKIAVSNCKTKAMTKVGTNEDM 195
DB 242 CHLGPWQIOLQKRMRYTGVIITOGAKRISPEYIKFYKIAVSNCKTKAMTKVGTNEDM 301
QY 196 VFRGNIDNTPYANSFTPIKAQYRLYPQVCRHCTLRMELLGCELSGCEPGLGMSGH 255
DB 302 VFRGNIDNTPYANSFTPIKAQYRLYPQVCRHCTLRMELLGCELSGCEPGLGMSGH 361
QY 256 IODVQITASSIFPTLNMDFMTPEPRKARLDKQKVNAMTSGHNDQSOMLQ----- 305
DB 362 IODVQITASSIFPTLNMDFMTPEPRKARLDKQKVNAMTSGHNDQSOMLQPTLLISVEVSG 421
QY 306 --VDLVPTRKYT 315
DB 422 TALALMVPTRAT 433

RESULT 12
US-60-207-315-523

; Sequence 523, Application us/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NOCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523
; LENGTH: 362
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-523

Query Match 62.1%; Score 1406.5; DB 23; Length 362;
Best Local Similarity 88.8%; Pred. No. 4,5e-122;
Matches 262; Conservative 3; Mismatches 7; Indels 23; Gaps 2;

QY 44 NINCEVEPCCKNGGICTDLVANYSCGCEPFGRNCOYKCSGPLGIEGGIISNOQITASS 103
DB 52 DINCEVEPCCKNGGICTDLVANYSCGCEPFGRNCOYKCSGPLGIEGGIISNOQITASS 111
QY 104 THRALFGLQKWYPYARLNKKGLINAMTAENDRWK-----RWIQLNLRKRMRY 152
DB 112 THRALFGLQKWYPYARLNKKGLINAMTAENDRWKIDELRGCHLGPMDQINLQKRMRY 171
QY 153 TGVITOGAKRISPEYIKFYKIAVSNCKTKAMTKVGTNEDMVFGRNIDNTPYANSFT 212
DB 172 TGVITOGAKRISPEYIKFYKIAVSNCKTKAMTKVGTNEDMVFGRNIDNTPYANSFT 231
QY 213 PPIKAQYRLYPQVCRHCTLRMELLGCELSGCEPGLGMSGHIDYQITASSIFPTLNM 272
DB 232 PPIKAQYRLYPQVCRHCTLRMELLGCELSGCEPGLGMSGHIDYQITASSIFPTLNM 291
QY 273 DMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQ-----VDLVPTRKYT 315
DB 292 DMFTPEPRKARLDKQKVNAMTSGHNDQSOMLQPTLLISVEVSGTALALMVPTRAT 346

RESULT 13

US-60-230-435-1519
; Sequence 1519, Application us/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NOCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1519
; LENGTH: 311
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(311)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-230-435-1519

Query Match 62.1%; Score 1405.5; DB 23; Length 311;
Best Local Similarity 89.1%; Pred. No. 4,6e-122;
Matches 262; Conservative 2; Mismatches 7; Indels 23; Gaps 2;

QY 45 INCEVEPCCKNGGICTDLVANYSCGCEPFGRNCOYKCSGPLGIEGGIISNOQITASS 104
|||||

Db 2 INEEVEPCNKGICTDVLVANSCECPGEFMRNCQYKCSGPLGIEGIIISNOQITASST 61
Qy 105 HRALEGLQKWPYYARLKNKGLINAMTAENDRMK-----RWIOINLORKMRYT 153
Db 62 HRALEGLQKWPYYARLKNKGLINAMTAENDRMWIOELRGCHGLPMDQINLORKMRYT 121
Qy 154 GVIOGAKRIGSPEYIKFYIAYSNQKTMWYKYGKTNEDMVPFGNIDNTPYANSFTP 213
Db 122 GVIOGAKRIGSPEYIKFYIAYSNQKTMWYKYGKTNEDMVPFGNIDNTPYANSFTP 181
Qy 214 PIKQOYVRLPYOVCRRHCTLMELLCCELSCGSEPLGKMSGHIDYOITASSIFRTLMND 273
Db 182 PIKQOYVRLPYOVCRRHCTLMELLCCELSCGSEPLGKMSGHIDYOITASSIFRTLMND 241
Qy 274 MFTWEPKARLDKQKYNAMTSGHNDOSQWLO-----VDLVPRTKVT 315
Db 242 MFTWEPKARLDKQKYNAMTSGHNDOSQWLOPLTISVSGTALAMVPIAT 295

RESULT 14
US-09-582-340-4
; Sequence 4, Application US/09582340
; GENERAL INFORMATION:
; APPLICANT: INSERM
; APPLICANT: INSITUUT CURIE
; APPLICANT: CNRS
; TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants
; TITLE OF INVENTION: thereof
; FILE REFERENCE: Lactadherin
; CURRENT APPLICATION NUMBER: US/09/582,340
; CURRENT FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 463
; TYPE: PRT
; ORGANISM: mammalian
US-09-582-340-4

Query Match 47.5%; Score 1075; DB 19; Length 463;
Best Local Similarity 50.1%; Pred. No. 3,8e-91;
Matches 203; Conservative 69; Mismatches 123; Indels 10; Gaps 6;
Qy 1 GPCPNCHNGSGTCEIS-EAYRGDTFIVGCKCPRGFNGIHQHNINCEVEPCNKGIC 59
Db 66 GPCPNCHNGSGTCEIS-EAYRGDTFIVGCKCPRGFNGIHQHNINCEVEPCNKGIC 122
Qy 60 TDLVANSCECPGEF--MGRNCQYKCSGPLGIEGIIISNOQITASSIFRTLMEL 117
Db 123 TTAVPNTAVPTPAFTPDLNLSRCSSTOLGMEGGLADSOISASVYVMGFMGLORNGPE 182
Qy 118 YARLKNKGLINAMTAENDRMKRWIOINLORKMRYTGVITOGAKRIGSPEYIKFYIAYSN 177
Db 183 LARLYKRGIVAMHASNYDS-KPMIOYNLIRKMRVSGVITOGAKRIGSPEYIKFYIAYSN 241
Qy 178 NDGKTMWYKYGKTNEDMVPFGNIDNTPYANSFTPPIKQOYVRLPYOVCRRHCTLMEL 237
Db 242 LDGRKFEFTIDE--SGGDKFELGMDNNSKLVNMFNPTLEAOYIRLYPVSCHRGCTLRFEEL 300
Qy 238 LGCELSCGSEPLGKMSGHIDYOITASSIFRTLMNDMFTWEPKARLDKQKYNAMTSGH 297
Db 301 LGCELHCLLEPLGLKNTTIPDSQMSASSKTYWNLRAFQWYRHLGRDLNCKKINAMTAOS 360
Qy 298 NDOSQWLOVDLVPTKYTGIIITOGAKRIGSPEYIKFYIAYSNDEHWTYODEKORRDK 357
Db 361 NSAKEMQLQVDLGTORQVYTGIIITOGAKRIGSPEYIKFYIAYSNDEHWTYODEKORRDK 418
Qy 358 VEOGNFNDTHRRKRVIDPPIYARHILPMSWYGRITLASELLGC 402
Db 419 VEOGNFNDTHRRKRVIDPPIYARHILPMSWYGRITLASELLGC 463

RESULT 15
US-09-582-340-2
; Sequence 2, Application US/09582340
; GENERAL INFORMATION:
; APPLICANT: INSERM
; APPLICANT: INSITUUT CURIE
; APPLICANT: CNRS
; TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants
; TITLE OF INVENTION: thereof
; FILE REFERENCE: Lactadherin
; CURRENT APPLICATION NUMBER: US/09/582,340
; CURRENT FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-582-340-2

Query Match 45.1%; Score 1020.5; DB 19; Length 387;
Best Local Similarity 48.8%; Pred. No. 3.4e-86;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;
Qy 3 CTPNCHNGSGTCEIS-EAYRGDTFIVGCKCPRGFNGIHQHNINCEVEPCNKGIC 61
Db 27 CTPNCHNGSGTCEIS-EAYRGDTFIVGCKCPRGFNGIHQHNINCEVEPCNKGIC 65
Qy 62 LVANSCECPGEFMRNCQYKCSGPLGIEGIIISNOQITASSIFRTLMEL 121
Db 66 -----CETKVEPLGEMNGINANSQIASSVVTPELGLHWPBELARL 108
Qy 122 NKGLINAMTAENDRMKRWIOINLORKMRYTGVITOGAKRIGSPEYIKFYIAYSN 181
Db 109 NKGLINAMTAENDRMKRWIOINLORKMRYTGVITOGAKRIGSPEYIKFYIAYSN 167
Qy 182 TW-AMRYKGTNEDMVPFGNIDNTPYANSFTPPIKQOYVRLPYOVCRRHCTLMEL 240
Db 168 EFDIHDVNRKHEFY--GNMKNVAVHMLFETPEAOYVRLPYOVCRRHCTLMEL 225
Qy 241 ELSCGSEPLGKMSGHIDYOITASSIFRTLMNDMFTWEPKARLDKQKYNAMTSGHND 300
Db 226 ELNCANPLGLKNSIPDKQITASSSYKTWGLHLPMSVYARLDKQKYNAMTSGHND 285
Qy 301 SOWLOVDLVPTKYTGIIITOGAKRIGSPEYIKFYIAYSNDEHWTYODEKORRDK 360
Db 286 SOWLOVDLVPTKYTGIIITOGAKRIGSPEYIKFYIAYSNDEHWTYODEKORRDK 345
Qy 361 GPNFNDTHRRKRVIDPPIYARHILPMSWYGRITLASELLGC 402
Db 346 GPNFNDTHRRKRVIDPPIYARHILPMSWYGRITLASELLGC 387

Search completed: May 23, 2001, 06:18:15
Job time: 409 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:55 ; Search time 28 Seconds
(without alignments)
103.938 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Perfect score: 2264
Sequence: 1 GPCPNPCNGHGTCEISEAV.....WSMYKRIITLASELGLCTERE 406

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41833 seqs, 7168127 residues

Total number of hits satisfying chosen parameters: 41833

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

Pending Patents AA New: *
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	453	20.0	931	US-09-583-638-4	Sequence 4, Appl
2	446.5	19.7	923	US-09-583-638-2	Sequence 2, Appl
3	446.5	19.7	957	PCT-US00-35017A-1285	Sequence 1285, Ap
4	234	10.3	734	PCT-US01-09226-65	Sequence 65, Appl
5	223	9.8	188	PCT-US01-01302-109	Sequence 109, Appl
6	206.5	9.1	723	US-09-828-366-21	Sequence 21, Appl
7	196.5	8.7	1964	US-09-467-997-1	Sequence 1, Appl
8	193.5	8.5	226	PCT-US01-01302-92	Sequence 92, Appl
9	188	8.3	904	PCT-US00-35017A-1079	Sequence 1079, Ap
10	187	8.3	383	PCT-US01-11988-770	Sequence 770, App
11	187	7.4	310	US-09-833-245-770	Sequence 770, App
12	166.5	7.3	201	PCT-US01-01302-103	Sequence 103, App
13	166	7.3	201	PCT-US01-10484-74	Sequence 74, Appl
14	156.5	6.9	213	PCT-US01-01302-90	Sequence 90, Appl
15	156.5	6.9	213	PCT-US01-01321-1559	Sequence 1559, Ap
16	148	6.5	3014	PCT-US01-010484-63	Sequence 63, Appl
17	143	6.3	1139	PCT-US01-01302-99	Sequence 99, Appl
18	143	6.3	1139	PCT-US01-01332-873	Sequence 873, Appl
19	141	6.2	4044	US-60-248-505-791	Sequence 791, App
20	140.5	6.2	379	US-09-808-689-4	Sequence 4, Appl
21	140.5	6.2	379	US-09-423-844-4	Sequence 4, Appl
22	140.5	6.2	379	US-09-403-296A-4	Sequence 4, Appl
23	140.5	6.2	379	US-09-380-139A-4	Sequence 4, Appl
24	137	6.1	122	PCT-US01-01302-97	Sequence 97, Appl
25	133.5	5.9	156	PCT-US01-01302-95	Sequence 95, Appl
26	132.5	5.9	4545	US-09-750-972-2	Sequence 2, Appl
27	125.5	5.5	501	PCT-US01-04098A-3346	Sequence 3346, Ap

28	125.5	5.5	576	1	PCT-US01-04098A-3820	Sequence 3820, Ap
29	125.5	5.5	576	1	PCT-US01-04098A-3821	Sequence 3821, Ap
30	125.5	5.5	603	1	PCT-US01-04098A-1853	Sequence 1853, Ap
31	125.5	5.5	650	1	PCT-US01-04098A-1852	Sequence 1852, Ap
32	125.5	5.5	4544	1	PCT-US01-04098A-1753	Sequence 1753, Ap
33	123.5	5.5	201	1	PCT-US01-01302-106	Sequence 106, Appl
34	123.5	5.5	534	5	US-09-804-156-14	Sequence 14, Appl
35	122	5.4	622	1	PCT-US01-04098A-3708	Sequence 3708, Ap
36	121	5.3	776	1	PCT-US01-04098A-1740	Sequence 1740, Ap
37	121	5.3	961	5	US-09-657-472-4	Sequence 4, Appl
38	119	5.3	415	5	US-09-802-704-7	Sequence 7, Appl
39	119	5.3	439	1	PCT-US01-11988-627	Sequence 627, App
40	119	5.3	439	1	PCT-US01-11988-628	Sequence 628, App
41	119	5.3	439	5	US-09-802-704-4	Sequence 4, Appl
42	119	5.3	439	5	US-09-833-245-627	Sequence 627, App
43	119	5.3	439	5	US-09-833-245-628	Sequence 628, App
44	118.5	5.2	509	5	US-09-423-844-315	Sequence 315, App
45	118.5	5.2	509	5	US-09-380-139A-314	Sequence 314, App

ALIGNMENTS

```
RESULT 1
US-09-583-638-4
; Sequence 4, Application US/09583638
; GENERAL INFORMATION:
; APPLICANT: KLAGSBRUN, MICHAEL
; APPLICANT: SOKER, SHAY
; APPLICANT: MAO, HUA-OUAN
; APPLICANT: TAKASHIMA, SEIJI
; TITLE OF INVENTION: NEUROPTILINS AND USE THEREOF IN METHODS FOR
; FILE REFERENCE: 701039-48800
; CURRENT APPLICATION NUMBER: US/09583, 638
; CURRENT FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US98/26127
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 60/069, 155
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 60/069, 687
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-583-638-4
```

Query Match 20.08; Score 453; DB 5; Length 931;
Best Local Similarity 35.44; Pred. No. 3.1e-35;
Matches 120; Conservative 59; Mismatches 122; Indels 38; Gaps 15;

```
QY 81 YKCSPLGIEGCIISNOITASSHRLALFLGOKWYPRYLANKGLINAWTAENDRMWR 140
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 275 FQCNVPLMEGSRGLANEISASST-----YDGRMTPOOSKLH--GDDNGMT--PNDSNNE 327
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 141 WIOINLQRMKRVTVGTTGA--KRIGSPEYIKFYKIVASNDGTWAMYKYKGTNEDMVR 198
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 328 YLVQVLRLLMTLMTALVAGALSRTONGYVYKSLKLEVTNGEDMYYR--HGRNH--KVFO 385
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 199 GNIDNNTPYANSFPPRIKAQYVRLPYQVCRHCHTLRRELLGCELSG--CSEPIGKMSGH 256
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 386 ANNDATFVNLKHALPLTRFVRIRPQTHSGIALRLLEFGCRVTDAPCSNMIGMLSGLI 445
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 257 GDYOTASSIFPTLMDQETWEPKARL--DKQ---KVNAWTSGHNDOSQMLQVLDLVP 311
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 446 AISOYSSAS-----TQELMSPSARLYVSSRSMFPRIRIQAPG-----EWMQVLDLGP 495
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 312 TVTGTITGAGNFGHV-----QFVGSYKLAYSNDGBHMTVTYDEKQKDKYFQGNFEN 365
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Db 496 KTYKGVIIIGAGGSGSIITAVARARFAVKKFVKSIVSLGKDWETIIDPRTIQPKLEFGNMHY 555

QY 366 DTHRKAVIDPPYIAKHILP-WSWYGRITLASELGC 402

Db 556 DTPDIRFD-PIPAQYVNVYPERKSPAG-IGMRLEVLGC 592

RESULT 2
US-09-583-638-2
: Sequence 2, Application US/09583638

APPLICANT: AKAASHKURU, MILHAEL
APPLICANT: SOKER, SHAY
APPLICANT: TAKASHIMA, SEIJI
APPLICANT: MIAO, HUA-OUAN
TITLE OF INVENTION: NEUROPEPTIDS AND USE T
TITLE OF INVENTION: DIAGNOSTICS AND PROGNO
FILE REFERENCE: 701039-48800
CURRENT APPLICATION NUMBER: US/09/583,638
CURRENT FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US98/26127
PRIOR FILING DATE: 1998-12-09
PRIOR APPLICATION NUMBER: 60/069,155
PRIOR FILING DATE: 1997-12-09
PRIOR APPLICATION NUMBER: 60/069,687
PRIOR FILING DATE: 1997-12-12
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 923
TYPE: PRT
ORGANISM: Homo sapiens
US-09-583-638-2

Query Match	19.7%;	Score 446.5;	DB 5;	Length 923;
Best Local Similarity	35.5%;	Pred. No. 1.3e-34;		
Matches 119;	Conservative 50;	Mismatches 135;	Indels 31	

[illegible]

```

RESULT      3
PCT-US00-35017A-1285
: Sequence 1285 Application PC/TUS0035017A
:
: GENERAL INFORMATION:
:
: APPLICANT: Hysq Inc
:
: TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
:
: FIL. REFERENCE: 784PCT
:
: CURRENT APPLICATION NUMBER: PCT/US00/35017A
:
: CURRENT FILING DATE: 2000-12-22

```

```

; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478

```

ORGANISM: Homo sapiens
PCT-US00-35017A-1285

Query Match	19.78;	Score 446.5;	DB 1;	Length 957;
Best Local Similarity	35.58;	Pred. No. 1.3e-34;		
Matches 119;	Conservative 50;	Mismatches 135;	Indels 31;	Gaps 12

[illegible]

```

RESULT 4
PCT-US01-09226-65
: Sequence 65, Application PC/TUS0109226
: GENERAL INFORMATION:
: APPLICANT: SMITHKLINE BEECHAM CORPORATION
: APPLICANT: SMITHKLINE BEECHAM P.L.C.
: TITLE OF INVENTION: NOVEL COMPOUNDS
: FILE REFERENCE: GPE0018
: CURRENT APPLICATION NUMBER: PCT/US01/09226
: CURRENT FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: 60/192,158
: PRIOR FILING DATE: 2000-03-24
: PRIOR APPLICATION NUMBER: 60/192,668
: PRIOR FILING DATE: 2000-03-27
: PRIOR APPLICATION NUMBER: 60/200,166
: PRIOR FILING DATE: 2000-04-27
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: PASTSEQ for Windows Version 3.0
: SEQ ID NO 65
: LENGTH: 734
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-US01-09226-65

```

```

Query Match:      10.3%   Score 234; DB 1: length 734;
Best Local Similarity 33.7%   Pred. No. 1,3e-14;
Matches 57; Conservative 25; Mismatches 71; Indels 16; Gaps 4.
OY 243 SCGSEPIGMKSGHIQDIQTASSIFRTLLNMDFTWEPKRLDKGV-----NNMTS 295
      :|: |||||::: |:: |||
Ob 113 TGC-PPLGTGLSTLVNSRLEASS-----SDFGLGPRGLNTQSLEDGEDLYDGAKCA 165

```


ORGANISM: Homo Sapien
US-09-828-366-21

Query Match
Best Local Similarity 42.3%; Pred. No. 5e-12;
Matches 33; Conservative 14; Mismatches 22; Indels 9; Gaps 2;

296 GANDQSMLOVDLIVPTKVTGIITOGAKDFGHOVFGSYKLANSNDGEHWYQDEKOR 355
166 EEOQADPWFQVDAGHPTBSGVITGGRNSVWRDYMTSYKQFSDNSRRTWSSRRHSSGM 225

356 DKVFGNFDNDTRKNVIPPPIYARHIRILPWSWY--GRITLASFLIGC 402
226 DAVEPANSDETPVNLPLPEPQVAFRILRLPOTWLOGGAPCLRAEILAC 274

RESULT 5
PCT-US01-01302-109
Sequence 109, Application PC/TUS0101302
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT207PCT
CURRENT APPLICATION NUMBER: PCT/US01/01302
CURRENT FILING DATE: 2001-01-17
NUMBER OF SEQ ID NOS: 192
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 109
LENGTH: 188
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (97)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (187)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (188)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-109

Query Match
Best Local Similarity 9.8%; Score 223; DB 1; Length 188;
Matches 42; Conservative 11; Mismatches 35; Indels 12; Gaps 3;

3 CTPNPGHNGTCEISFAYRGDFIIGYCKPRGFNGIHCHQHNINECEVEPRCKNGICTD 61
78 CDSNPCRNGSGCKDQED-----GXCCPFGYGLCHESHTSCADSPCFNGSGCRER 130

62 -LVANYSCECPGEPFGRNCQY---CSGPLIGEGIIISNQ 97
131 NQGANVACECPNPTGSCNCEKVDCTSNPCANGGQCLNR 170

RESULT 6
US-09-828-366-21
Sequence 21, Application US/09828366
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Klein, Robert D.
APPLICANT: Napier, Mary
APPLICANT: Wood, William I.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
TITLE OF INVENTION: CELL GROWTH
FILE REFERENCE: P1694R1C1
CURRENT APPLICATION NUMBER: US/09/828,366
CURRENT FILING DATE: 2001-04-05
Prior filing data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 29
SEQ ID NO 21
LENGTH: 723
TYPE: PRT

ORGANISM: Homo Sapien
US-09-828-366-21

Query Match
Best Local Similarity 9.1%; Score 206.5; DB 5; Length 723;
Matches 33; Conservative 14; Mismatches 22; Indels 9; Gaps 2;

3 CTPNPGHNGTCEISFAYRGDFIIGYCKPRGFNGIHCHQHNINECEVEPRCKNGICTD 61
409 CSSPFCSSNGAKCYDLGDA-----YLCRCQAGFSGRHCDNDVDCASSPCANGSGICTD 460

62 LVANYSCECPGEPFGRNC 79
461 GVNDPSCPCPGYTGRCN 478

RESULT 7
US-09-467-997-1
Sequence 1, Application US/09467997
GENERAL INFORMATION:
APPLICANT: Kitajewski, Jan
TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
FILE REFERENCE: 53863-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/467,997
CURRENT FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1964
TYPE: PRT
ORGANISM: mouse
US-09-467-997-1

Query Match
Best Local Similarity 8.7%; Score 196.5; DB 5; Length 1964;
Matches 31; Conservative 12; Mismatches 29; Indels 7; Gaps 1;

2 CTPNPGHNGTCEISFAYRGDFIIGYCKPRGFNGIHCHQHNINECEVEPRCKNGICTD 61
929 PCENPCHNGSTCYPPPS-----GYVQCAPGIEGONCSKVIDACOSPCNHRGTCTIS 981

62 LVANYSCECPGEPFGRNCQ 80
982 RFGGFHCACPPGFVGLRCE 1000

RESULT 8
PCT-US01-01302-92
Sequence 92, Application PC/TUS0101302
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT207PCT
CURRENT APPLICATION NUMBER: PCT/US01/01302
CURRENT FILING DATE: 2001-01-17
NUMBER OF SEQ ID NOS: 192
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 92
LENGTH: 226
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (133)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-92

Query Match
Best Local Similarity 8.5%; Score 193.5; DB 1; Length 226;
Matches 29.7%; Pred. No. 1.9e-11;

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 770
; LENGTH: 383
```

ORGANISM: Homo sapiens
PCT-US01-11988-770

Query Match	8.3%	Score 187;	DB 1;	Length 383;
Normalized Distance	38.2%	Need 1	Cost 10	

Matches 31; Conservative 14; Mismatches 30; Indels 6; Gaps 2.

0y 1 GPC--TPNCHNGSTCEISEAYRGDTFIGYVCKPRGFNGIHCQHININECEVEPECKNGGI 58

QY 59 CTDLVANSCECPGEFMGRNC 79
| | : || | | | |
100 GRTATWGGGCGTGTATTGGG 200

RECITE 11

US-09-833-245-770

; GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences Inc

FILE REFERENCE: PF546BCT

CURRENT FILING DATE: 2001-04-12

PRIOR FILING DATE: 2000-04-12

PRIOR FILING DATE: 2000-12-21

PRIOR FILING DATE: 2000-04-25

; SOFTWARE: PatentIn Ver. 2.1

```

;      TYPE: PRT
;      ORGANISM: Homo sapiens
US-09-833-245-770

```

0
Z
-
L

0
0
0

0

A
0
1

9
J
n

,

-
-

0
0
0

Best Local Similarity 38.38; Pred. No. 1.6e-10;

1 000 ENU/ML

```

| | : | | | : : : | : | | | : | : | : | | |
133 CCGCAGCCGACGGGCCCGGCGT - - T AATGCGGT VCEVCA GCEUUNDDCT WPDCAVCAT 180

```

00 50 CREDIT ADVANCE 70

```

189 CTGCTNPFSCCTCPECEACBPCC 200
      | | : | | | | | |

```

RESULT 12

Sequence 103, Application PC/TUS0101302

APPLICANT: Human Genome Sciences, Inc., et al.

FILE REFERENCE: PT207PCT

CURRENT FILING DATE: 2001-01-17

; SOFTWARE; PatentIn Ver. 2.0
DEPT NO 103

LENGTH: 310

ORGANISM: Homo sapiens

PCT-US01-01302-103

Query Match 7.4%; Score 166.5; DB 1; Length 310;
Best Local Similarity 42.7%; Pred. No. 1e-08;
Matches 35; Conservative 7; Mismatches 31; Indels 9; Gaps 5;

OY 7 PCHNGGCEISEAVRGDTFFIGYVCKPRGFGNGIHCOHINCEVEPCCKNGGICD--DLVA 64
DB 69 PCVNEGMC---VYIHNGT--GI-CKCPBGFLBETCOHR-DPECKNRCONGGTCVAQAMLG 121
OY 65 NYSCCEPGEFMRNCOYKCSGP 86
DB 122 KATCRCAAGFTGEDCOYSTSHP 143

RESULT 13

PCT-US01-10484-74

Sequence 74, Application PC/TUS0110484

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc

TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 21272-048

CURRENT APPLICATION NUMBER: PCT/US01/10484

PRIOR FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: 09/540,217

PRIOR FILING DATE: 2000-03-31

PRIOR APPLICATION NUMBER: 09/649,167

PRIOR FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: 09/668,680

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: 09/695,618

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 09/728,711

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: NOT YET ASSIGNED

PRIOR FILING DATE: 2000-03-14

NUMBER OF SEQ ID NOS: 172

SOFTWARE: Custom

SEQ ID NO 74

LENGTH: 201

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US01-10484-74

Query Match 7.3%; Score 166; DB 1; Length 201;
Best Local Similarity 33.6%; Pred. No. 6.6e-09;
Matches 38; Conservative 24; Mismatches 43; Indels 8; Gaps 3;

OY 279 PRKARLDKQGVNMTSGHNDQOWLOVDLVTYKVTGIITGCAKDFHGVFGSYKLA 338
DB 56 POFSLNRDQAGWIPVSNKTYQIDLDGERMEYTAATGG--YSSDQWVSYLLMF 113
OY 339 SNDGHEWTVYODEKORDKY--FOGNFNDTHRKNVIDPPYARHRIILPSMW 389
DB 114 SDGGRNMKQY---RREESIMGPPGNTNADSVVHYRLQPPFARLRPLAW 162

RESULT 14

PCT-US01-01302-90

Sequence 90, Application PC/TUS0101302

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc., et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT207PCT

CURRENT APPLICATION NUMBER: PCT/US01/01302

PRIOR FILING DATE: 2001-01-17

NUMBER OF SEQ ID NOS: 192

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 90

LENGTH: 213

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (1)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (17)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (179)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (195)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (205)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

PCT-US01-01302-90

Query Match 6.9%; Score 156.5; DB 1; Length 213;
Best Local Similarity 35.4%; Pred. No. 5.7e-08;
Matches 28; Conservative 10; Mismatches 36; Indels 5; Gaps 1;

OY 2 PCTPNCHNGTCEISEAVRGDTFFIGYVCKPRGFGNGIHCOHINCEVEPCCKNGGICD 61
DB 98 PCFSNPGGSAITCVNTPGER-----SFLCKPBGISGTICETTGSCGKNSCQHGICHQ 152
OY 62 LVANYSCEPGEFMRNCO 80
DB 153 DPYVPCICPAGYAGRFCE 171

RESULT 15

PCT-US01-01321-1559

Sequence 1559, Application PC/TUS0101321

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc., et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC011PCT

CURRENT APPLICATION NUMBER: PCT/US01/01321

PRIOR FILING DATE: 2001-01-17

NUMBER OF SEQ ID NOS: 2181

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1559

LENGTH: 213

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (1)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (17)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (179)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (195)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (205)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

PCT-US01-01321-1559

Query Match 6.9%; Score 156.5; DB 1; Length 213;
Best Local Similarity 35.4%; Pred. No. 5.7e-08;
Matches 28; Conservative 10; Mismatches 36; Indels 5; Gaps 1;

OY 2 PCTPNCHNGTCEISEAVRGDTFFIGYVCKPRGFGNGIHCOHINCEVEPCCKNGGICD 61
DB 98 PCFSNPGGSAITCVNTPGER-----SFLCKPBGISGTICETTGSCGKNSCQHGICHQ 152
OY 62 LVANYSCEPGEFMRNCO 80
DB 153 DPYVPCICPAGYAGRFCE 171

Wed May 23 06:29:17 2001

us-09-237-981-14_copy_108_513.rapn

Page 6

Db 98 PCFSNPGQSGATCVNTPGER-----SFLCKCPGYSCTICETITGSCGKNSCOHGICHQ 152
QY 62 LVANYSCGPGEFMRNQ 80
| 11 : 11 1:
Db 153 DRIYPCICPAGYAGRFE 171

Search completed: May 23, 2001, 06:13:56
Job time: 150 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:21 ; Search time 38.64 Seconds

(without alignments)
201.853 Million cell updates/sec

Title: US-09-237-981-14_COPY_108_513

Perfect score: 2264
Sequence: 1 GPCPNPCHNGSTCEISRAY.....MSWVGRTLASLGLCTERE 406

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2264	100.0	513	2	US-08-480-229C-14
2	2264	100.0	513	2	US-08-659-235C-14
3	2170.5	95.9	480	2	US-08-480-229C-10
4	2170.5	95.9	480	2	US-08-659-235C-10
5	1717	75.8	321	2	US-08-480-229C-21
6	1717	75.8	321	2	US-08-659-235C-21
7	1063	47.0	463	2	US-08-162-402B-9
8	1063	47.0	463	2	US-08-162-402B-8
9	1020.5	45.1	387	2	US-08-162-402B-6
10	948	41.9	320	2	US-08-480-229C-20
11	948	41.9	320	2	US-08-659-235C-20
12	784.5	34.7	221	2	US-08-480-229C-29
13	784.5	34.7	221	2	US-08-659-235C-29
14	682.5	30.1	2183	3	US-08-746-111-5
15	670	28.6	2319	1	US-08-212-133A-8
16	670	28.6	2319	1	US-08-474-503-6
17	670	28.6	2319	2	US-08-670-707A-6
18	670	28.6	2319	4	US-09-037-601-6
19	670	28.6	2319	5	PCT-US94-13200-6
20	653	28.8	2332	1	US-08-216-594A-2
21	650	28.7	2351	6	5422260-1
22	649	28.7	1471	1	US-08-683-839B-3
23	649	28.7	1661	2	US-08-882-083-2
24	649	28.7	1661	2	US-08-558-107-2
25	649	28.7	1661	4	US-09-243-539-2
26	649	28.7	2332	1	US-07-864-004B-4
27	649	28.7	2332	1	US-08-251-937A-4

28	649	28.7	2332	1	US-08-212-133A-2	Sequence 2, Appl1
29	649	28.7	2332	1	US-08-474-503-2	Sequence 2, Appl1
30	649	28.7	2332	2	US-08-670-707A-2	Sequence 2, Appl1
31	649	28.7	2332	4	US-09-037-601-2	Sequence 2, Appl1
32	649	28.7	2332	5	PCT-US93-03275-4	Sequence 2, Appl1
33	649	28.7	2332	5	PCT-US94-13200-2	Sequence 2, Appl1
34	649	28.7	2351	1	US-08-121-202-2	Sequence 2, Appl1
35	649	28.7	2351	1	US-08-366-851A-2	Sequence 2, Appl1
36	643	28.4	217	1	US-07-607-538C-3	Sequence 3, Appl1
37	643	28.4	217	2	US-08-162-402B-3	Sequence 3, Appl1
38	643	28.4	218	1	US-07-607-538C-2	Sequence 2, Appl1
39	643	28.4	218	2	US-08-162-402B-2	Sequence 2, Appl1
40	643	28.2	2351	6	5171844-2	Sequence 2, Appl1
41	643	28.2	1443	2	US-08-670-707A-39	Sequence 39, Appl1
42	639	28.2	1443	4	US-09-037-601-39	Sequence 39, Appl1
43	639	28.2	2133	2	US-08-670-707A-37	Sequence 37, Appl1
44	639	28.2	2133	4	US-09-037-601-37	Sequence 37, Appl1
45	501	22.1	157	2	US-08-162-402B-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-08-480-229C-14
; Sequence 14, Application US/08480229C
; Patent No. 5674562
; GENERAL INFORMATION:
; APPLICANT: Quentromous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 26,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-480-229C-14

Query Match 100.0%; Score 2264; DB 2; Length 513;
Best local similarity 100.0%; Pred. No. 5,8e+180;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1  GPCTPNCHNGGTCISEAVRGDTFIGYCKCPRGFNHICQHNINECEVEPCKNNGICT 60
          |||||||
Db      108 GPCTPNCHNGGTCISEAVRGDTFIGYCKCPRGFNHICQHNINECEVEPCKNNGICT 167
QY      61  DLVANYSCGCEPGFEMGRNCOYKCSGFLGIEGGIISNOQITASSTHRALFGLQWPPYYAR 120
          |||||||
Db      168 DLVANYSCGCEPGFEMGRNCOYKCSGFLGIEGGIISNOQITASSTHRALFGLQWPPYYAR 227
QY      121 LNKGLINAMTAENDRMKRWIOINLQRMRYGVITOGAKRIGSPYIKFYKIASNDG 180
          |||||||
Db      228 LNKGLINAMTAENDRMKRWIOINLQRMRYGVITOGAKRIGSPYIKFYKIASNDG 287
QY      181 KTWAMYKVGKTNEDMYFRGNIDNTPYANSFTPPKAOYVLYPQVCRHCTLRMELLCG 240
          |||||||
Db      288 KTWAMYKVGKTNEDMYFRGNIDNTPYANSFTPPKAOYVLYPQVCRHCTLRMELLCG 347
QY      241 ELSCGSEPLGKSGHIDQVITASSIFRTLAMDFTWEPKARLDKQKVNAMTSGHNDQ 300
          |||||||
Db      348 ELSCGSEPLGKSGHIDQVITASSIFRTLAMDFTWEPKARLDKQKVNAMTSGHNDQ 407
QY      301 SOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHWTYODEKORKDYVQ 360
          |||||||
Db      408 SOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHWTYODEKORKDYVQ 467
QY      361 GNFDNDTHRKNVIDPPIYAHIRILPMSWYGRITLASLGLCTEE 406
          |||||||
Db      468 GNFDNDTHRKNVIDPPIYAHIRILPMSWYGRITLASLGLCTEE 513

```

```

RESULT      2
US-08-659-235C-14
; Sequence 14, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-659-235C-14

```

Query Match 100.0%; Score 2264; DB 2; Length 513;
 Best Local Similarity 100.0%; Pred. No. 5 Be-180;
 Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1  GPCTPNCHNGGTCISEAVRGDTFIGYCKCPRGFNHICQHNINECEVEPCKNNGICT 60
          |||||||
Db      108 GPCTPNCHNGGTCISEAVRGDTFIGYCKCPRGFNHICQHNINECEVEPCKNNGICT 167
QY      61  DLVANYSCGCEPGFEMGRNCOYKCSGFLGIEGGIISNOQITASSTHRALFGLQWPPYYAR 120
          |||||||
Db      168 DLVANYSCGCEPGFEMGRNCOYKCSGFLGIEGGIISNOQITASSTHRALFGLQWPPYYAR 227
QY      121 LNKGLINAMTAENDRMKRWIOINLQRMRYGVITOGAKRIGSPYIKFYKIASNDG 180
          |||||||
Db      228 LNKGLINAMTAENDRMKRWIOINLQRMRYGVITOGAKRIGSPYIKFYKIASNDG 287
QY      181 KTWAMYKVGKTNEDMYFRGNIDNTPYANSFTPPKAOYVLYPQVCRHCTLRMELLCG 240
          |||||||
Db      288 KTWAMYKVGKTNEDMYFRGNIDNTPYANSFTPPKAOYVLYPQVCRHCTLRMELLCG 347
QY      241 ELSCGSEPLGKSGHIDQVITASSIFRTLAMDFTWEPKARLDKQKVNAMTSGHNDQ 300
          |||||||
Db      348 ELSCGSEPLGKSGHIDQVITASSIFRTLAMDFTWEPKARLDKQKVNAMTSGHNDQ 407
QY      301 SOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHWTYODEKORKDYVQ 360
          |||||||
Db      408 SOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHWTYODEKORKDYVQ 467
QY      361 GNFDNDTHRKNVIDPPIYAHIRILPMSWYGRITLASLGLCTEE 406
          |||||||
Db      468 GNFDNDTHRKNVIDPPIYAHIRILPMSWYGRITLASLGLCTEE 513

```

```

RESULT      3
US-08-480-229C-10
; Sequence 10, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 480 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-229C-10

Query Match 95.9%; Score 2170.5; DB 2; Length 480;
 Best Local Similarity 95.8%; Pred. No. 3e-172;
 Matches 389; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 1 GCCTPNCNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHINCEVEPCNGSIGCT 60
 DB 76 GCCTPNCNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHINCEVEPCNGSIGCT 135
 QY 61 DLVANYSCCEGPEFMGRNCOYKCSGPLGIEGGIISNOQITASSSTRALFGLQKWPYYAR 120
 DB 136 DLVANYSCCEGPEFMGRNCOYKCSGPLGIEGGIISNOQITASSSTRALFGLQKWPYYAR 195
 QY 121 LNKKGILNMTAENDRWKRWIOINLQKRWYGVITOGAKRIGSPEYIKFYKIAYSNDG 180
 DB 196 LNKKGILNMTAENDRWKRWIOINLQKRWYGVITOGAKRIGSPEYIKFYKIAYSNDG 254
 QY 181 KTWAMYKVKGTNEEDVFRGNIDNNTPYANSFPPIKAOYVRLYPQVCRHCTLMELLGC 240
 DB 255 KTWAMYKVKGTNEEDVFRGNIDNNTPYANSFPPIKAOYVRLYPQVCRHCTLMELLGC 314
 QY 241 ELSCGSEPLGKSGHIDYOITASSIFRTLNDMTWEPKARLDKQKVAWMTSGHNDQ 300
 DB 315 ELSCGSEPLGKSGHIDYOITASSIFRTLNDMTWEPKARLDKQKVAWMTSGHNDQ 374
 QY 301 SOWLOYDLVPTKYTGIIITOGAKDFGVQSVSYKIAYSNDGEHMTYQDEKOKKRVFO 360
 DB 375 SOWLOYDLVPTKYTGIIITOGAKDFGVQSVSYKIAYSNDGEHMTYQDEKOKKRVFO 434
 QY 361 GNFDNTHRKKNVIDPPIYARHRIILPMSWGRITLASELLGCTEE 406
 DB 435 GNFDNTHRKKNVIDPPIYARHRIILPMSWGRITLASELLGCAEE 480

RESULT . 4

US-08-659-235C-10
 ; Sequence 10, Application US/08659235C
 ; Patent No. 3677281
 ; GENERAL INFORMATION:
 ; APPLICANT: Quentermous, Thomas
 ; APPLICANT: Hogan, Bridgid
 ; APPLICANT: Snodgrass, H. Ralph
 ; APPLICANT: Zupancic, Thomas J.
 ; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 ; TITLE OF INVENTION: CELL LOCUS-1
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/659,235C
 ; FILING DATE: 05-JUN-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Poissant, Brian M.
 ; REGISTRATION NUMBER: 28,462
 ; REFERENCE/DOCKET NUMBER: 8907-0034-999
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELE: 66141 Pennie
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 480 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-659-235C-10

Query Match 95.9%; Score 2170.5; DB 2; Length 480;
 Best Local Similarity 95.8%; Pred. No. 3e-172;
 Matches 389; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 1 GCCTPNCNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHINCEVEPCNGSIGCT 60
 DB 76 GCCTPNCNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHINCEVEPCNGSIGCT 135
 QY 61 DLVANYSCCEGPEFMGRNCOYKCSGPLGIEGGIISNOQITASSSTRALFGLQKWPYYAR 120
 DB 136 DLVANYSCCEGPEFMGRNCOYKCSGPLGIEGGIISNOQITASSSTRALFGLQKWPYYAR 195
 QY 121 LNKKGILNMTAENDRWKRWIOINLQKRWYGVITOGAKRIGSPEYIKFYKIAYSNDG 180
 DB 196 LNKKGILNMTAENDRWKRWIOINLQKRWYGVITOGAKRIGSPEYIKFYKIAYSNDG 254
 QY 181 KTWAMYKVKGTNEEDVFRGNIDNNTPYANSFPPIKAOYVRLYPQVCRHCTLMELLGC 240
 DB 255 KTWAMYKVKGTNEEDVFRGNIDNNTPYANSFPPIKAOYVRLYPQVCRHCTLMELLGC 314
 QY 241 ELSCGSEPLGKSGHIDYOITASSIFRTLNDMTWEPKARLDKQKVAWMTSGHNDQ 300
 DB 315 ELSCGSEPLGKSGHIDYOITASSIFRTLNDMTWEPKARLDKQKVAWMTSGHNDQ 374
 QY 301 SOWLOYDLVPTKYTGIIITOGAKDFGVQSVSYKIAYSNDGEHMTYQDEKOKKRVFO 360
 DB 375 SOWLOYDLVPTKYTGIIITOGAKDFGVQSVSYKIAYSNDGEHMTYQDEKOKKRVFO 434
 QY 361 GNFDNTHRKKNVIDPPIYARHRIILPMSWGRITLASELLGCTEE 406
 DB 435 GNFDNTHRKKNVIDPPIYARHRIILPMSWGRITLASELLGCAEE 480

RESULT . 5

US-08-480-229C-21
 ; Sequence 21, Application US/08480229C
 ; Patent No. 5874562
 ; GENERAL INFORMATION:
 ; APPLICANT: Quentermous, Thomas
 ; APPLICANT: Hogan, Bridgid
 ; APPLICANT: Snodgrass, H. Ralph
 ; APPLICANT: Zupancic, Thomas J.
 ; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 ; TITLE OF INVENTION: CELL LOCUS-1
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,229C
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536

1 TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
2 TITLE OF INVENTION: GLOBULE (HMF) ANTIGEN, FRAGMENTS & FUSION PROTEIN
3 NUMBER OF SEQUENCES: 29
4 CORRESPONDENCE ADDRESS:
5 ADDRESSEE: Pretty, Schroeder & Poplawski
6 STREET: 444 South Flower St., 19th Floor
7 CITY: Los Angeles
8 STATE: CA
9 COUNTRY: USA
10 ZIP: 90071
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Diskette
13 COMPUTER: IBM Compatible
14 OPERATING SYSTEM: DOS
15 SOFTWARE: FastSeq for Windows Version 2.0
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/162,402B
18 FILING DATE: 03-DEC-1993
19 CLASSIFICATION: 435
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER:
22 FILING DATE:
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Amzel, Viviana
25 REGISTRATION NUMBER: 30,930
26 REFERENCE/DOCKET NUMBER: P66 38215
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 213-622-7700
29 TELEFAX: 213-489-4210
30 TELEX:
31 INFORMATION FOR SEQ ID NO: 6:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 387 amino acids
34 TYPE: amino acid
35 STRANDEDNESS: unknown
36 TOPOLOGY: unknown
37 MOLECULE TYPE: peptide
38 US-08-162-402B-6

Query Match 45.1%; Score 1020.5; DB 2; Length 387;
Best Local Similarity 48.8%; Pred. No. 4.6e-77;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;
1 3 CTPNCHNGGTC-EISEATYAGDTFVIGYCKCPRGENGHCQHINECEVEPECKNGICTD 61
2 27 CSKNCCHNGGTC-EISEATYAGDTFVIGYCKCPRGENGHCQHINECEVEPECKNGICTD 61
3 62 LVANVSCCEPGEFGRNCOYKCSGPIGEGGIIISNOQITASSTHRALFGLQKWPYARL 121
4 66 -----CEIKVEPIELMENGNIANSQIASSVRYTFGLQHWPELARKL 108
5 122 NKKGLINAMTAENDRNKRMIOINLQKMKVTVITOGAKRIGSPYIKFYKIAVSDNGK 181
6 109 NRAGVNVNAMPSSND-NPMIOVNLRLRMVTVGTVOGASRLASHEYLKAFKVAATSLNGH 167
7 182 TW-AMVYKKGNEDEVRGINDNNTPYANSFPPRIKAQVVRVLPYOVCRHCHTLRELLGC 240
8 168 EEDFDIHNKKHKEFY--GMMNKNAAVHVFETPRVAQVVRVLPYOVCRHCHTLRELLGC 225
9 241 ELSCGSEPLKMGKSHIDYQVITASSIFRILNMDFTWEPKARLDKOGKVNAMTSGHNDQ 300
10 226 ELNCGANLGLKNSIDKQITASSSYKTWGLHLSMNSYARLDKOGKFNAMVAGSYGN 285
11 301 SOMLOVDLVPYTKYGTGIIITOGADFGHVOFVSGKLAYSNDGEHMTVYODEKQKRDYFQ 360
12 286 DQMLQVDLGSSEKVTGIIITOGARNFSGVQFVASYKVAIYNSDSANMTEYQDPRTGSKLTFP 345
13 361 GNFDNDTHRKAVIDPPIYARHRIELFMSWYGRITLASSELLGC 402
14 346 GMDVHSHKKNLFETPIIARYVRILPVAMHNRITALELLGC 387

RESULT 10
US-08-480-229C-20
1 Sequence 20; Application US/08480229C
2 Patent No. 3874562
3 GENERAL INFORMATION:
4 APPLICANT: Quarterman, Thomas
5 APPLICANT: Hogan, Bridgid
6 APPLICANT: Snodgrass, H. Ralph
7 APPLICANT: Zupancic, Thomas J.
8 TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
9 NUMBER OF SEQUENCES: 29
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Pennile & Edmonds LLP
12 STREET: 1155 Avenue of the Americas
13 CITY: New York
14 STATE: New York
15 COUNTRY: United States
16 ZIP: 10036-2711
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: PatentIn Release #1.0, Version #1.30
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/480,229C
24 FILING DATE: 07-JUN-1995
25 CLASSIFICATION: 536
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Poissant, Brian M.
28 REGISTRATION NUMBER: 28,462
29 REFERENCE/DOCKET NUMBER: 8907-0026-999
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (212) 790-9090
32 TELEFAX: (212) 869-8864/9741
33 TELEX: 66141 pennile
34 INFORMATION FOR SEQ ID NO: 20:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 320 amino acids
37 TYPE: amino acid
38 STRANDEDNESS:
39 TOPOLOGY: unknown
40 MOLECULE TYPE: protein
41 US-08-480-229C-20

Query Match 41.9%; Score 948; DB 2; Length 320;
Best Local Similarity 54.7%; Pred. No. 3.7e-71;
Matches 175; Conservative 50; Mismatches 95; Indels 0; Gaps 0;
1 83 CGSPGIGSGGIISNOQITASSTHRALFGLQKWPYARLKKGLINAMTAENDRNKRM 142
2 1 CSTQGMGGAIAQSISASVYVGMGLQMGPELARIYRGIIVAMHNASVYDSKPMI 60
3 143 QNLQKRRKRVGVITOGAKRIGSPYIKFYKIAVSDNGKTMAMKYKGTNEDMVRGND 202
4 61 QVNLKRRKRVSGVMTQGSRAKREYLKTFKVAISLDGAKKEFIQDESGGKFEGLND 120
5 203 NNTPIANSTFPIKAQVVRVLPYOVCRHCHTLRELLGCSEPLKMGKSHIDYQVIT 262
6 121 NNSLVNMFNPLEAOYRLPVSCARGCTLRELLGCSEPLKMGKSHIDYQVIT 180
7 263 ASSIFRILNMDFTWEPKARLDKOGKVNAMTSGHNDQSOMLQVDLVPYTKYGTGII 322
8 181 ASSSYKTNLNAIFGYPHILGSLDNGKINAMTQNSAKEMLOVDLGTQROYTGII 240
9 323 KQFGHVOFVSGKLAYSNDGEHMTVYODEKQKRDYFQGNFNDTHRKAVIDPPI 382
10 241 RQFGHIOFVSGKVAHSDGVQMTVYXXBQSSKVFQGLNNSHKKNIFKRPMA 300
11 383 RILPMSWYGRITLASSELLGC 402
12 301 RVLPSVWNRITLRELLGC 320

COMPUTER READABLE FORM:

;;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2319 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; AUTHORS: Elder, F.
; AUTHORS: Gitschler, J.
; TITLE: Sequence of the Murine Factor VIII cDNA.
; Patent No. 5663060
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 8: FROM 1 TO 2319
; US-08-212-133A-8

Query Match 29.6%; Score 670; DB 1; Length 2319;

Best Local Similarity 46.3%; Pred. No. 4.8e-47;

Matches 151; Conservative 40; Mismatches 119; Indels 16; Gaps 5;

QY 82 KCGSPGIEGGIISNOQITASSSTHRAFLFGLOKWPYARLNKKGLINAWTAENDRKRW 141
DB 2007 QCOIPGAMSGSIRDFQITNSGHY-----GQWAPNLARLHSGSINAWSTKEP---FSW 2057
QY 142 IQINLOKRMVTVITOGAKRISPEYIKFYKAYVNDGKTAMKVKGTINEDMVERGNI 201
DB 2058 IKYDILAPMIVHGIKTGAKQKFSLSYISQFILMSLDGKKWLSYGNSGTGLMVEFGNV 2117
QY 202 DNNTPYANSTFPRIKAOYVLYLPQVCRHCTLMELICGLSCGSEPLGKMSGHIDYQI 261
DB 2118 DSSGIKINSFNPRIARVIRLHPHSHSIRSTLMELMGCDLNSCSIPLGMSKVISDTQI 2177
QY 262 TASSIFPTLMDMF-TWEPKARLDKOGKVNAMTSGHNDQSOQLQVLDLVPYTVGIIITQ 320
DB 2178 TASSYF-----TNPATWSPSQARLHLQGRINARPOVNDPKQWLDQKTKAKVIGIITQ 2233
QY 321 GANDFGHVCDFVGSYKLAAYVNDGSHWTVYODEKORCKDVFGQNFNDNTHRKKNVIDPPIYAR 380
DB 2234 GVKSLFTSMVYKVELLSSQDGHMT--QILYNGKVKVFGQNGDSSSTPMNNSLDPLLR 2291
QY 381 HIRLPMSWTGRITLASSELLGTEEE 406
DB 2292 YLRHPQIWEHQAIRLEIIGCEAAQ 2317

Search completed: May 23, 2001, 06:13:23
Job time: 117 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:12:39 ; Search time 62.53 Seconds
(without alignments)
528.638 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513

Perfect score: 2675
Sequence: 1 MKRSVAVMLVGLSLGVPQF.....MSWYGRITLASLGLCTEEE 481

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 67: *
1: pirl: *
2: pirl: *
3: pirl: *
4: pirl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1173	43.9	409	2	T11743
2	1156	43.2	463	1	A36479
3	1154	43.1	427	2	JC4915
4	1145.5	42.8	427	2	S74211
5	1119	41.8	401	2	S65138
6	682.5	25.5	2183	2	T42764
7	670.5	25.1	2224	1	KFH05
8	670	25.0	2319	2	A47004
9	657.5	24.6	2211	1	KFB05
10	649	24.3	2351	1	EZHU
11	643	24.0	218	2	A47285
12	639	23.9	2133	2	T42763
13	464.5	17.4	216	2	A44258
14	450	16.8	927	1	JQ0948
15	374	14.0	3133	2	S52093
16	330	12.3	2352	2	T30201
17	326	12.2	1064	2	A40136
18	313.5	11.7	2524	2	A35844
19	312	11.7	2531	2	A18188
20	311.5	11.6	2531	2	A46019
21	307	11.5	2437	2	S42612
22	303.5	11.3	570	2	A48836
23	302.5	11.3	2703	1	A24420
24	297	11.1	1203	2	A49175
25	297	11.1	2531	2	T31070
26	296	11.1	2471	2	A49128
27	292.5	10.9	1408	2	S16148
28	291.5	10.9	2318	2	S45306

30	288.5	10.8	2321	2	S78549	notch3 protein - h
31	285.5	10.7	728	2	I50719	C-Delta-1 - chicke
32	283	10.6	1531	2	T42218	slit-1 protein hom
33	282.5	10.6	861	2	A48825	Notch homolog Motc
34	274.5	10.3	832	2	A31246	neurogenic protein
35	274.5	10.3	833	2	S19087	gene Delta protein
36	274.5	10.3	880	2	S00670	neurogenic repeat
37	274.5	10.3	1687	2	T30176	EGF repeat transme
38	271.5	10.1	1220	2	A56136	jagged protein pre
39	268	10.0	473	2	A56175	adhesive plaque pr
40	267	10.0	722	2	I48324	DELTA-like 1 - mou
41	262.5	9.8	845	2	JC5256	adipocyte transcri
42	262.5	9.8	2139	2	A35672	crumbs protein - f
43	259.5	9.7	1469	2	B36665	slit protein 2 pre
44	259.5	9.7	1480	2	A36665	slit protein 1 pre
45	259	9.7	1964	2	T09059	notch4 - mouse

ALIGNMENTS

RESULT 1	T11743	PP47 protein - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)		
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000		
C:Accession: T11743		
R:Ensslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toe Biol. Reprod. 58, 1057-1064, 1998		
A:Title: Molecular cloning and characterization of P47, a novel boar sperm-associated A:Reference number: 217325; MUID:98206817		
A:Accession: T11743		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: mRNA		
A:Residues: 1-409 <ENS>		
A:Cross-References: EMBL:Y11683; NID:q2652927; PIDN:CAA72379.1; PID:q2652928		
A:Experimental source: testis		
C:Function:		
A:Description: may be involved in membrane remodeling and/or function as a zona pellu C:Superfamily: unassigned EGF-related proteins; EGF homology		
F:6-40/Domain: EGF homology <EGF>		
Query Match	43.9%;	Score 1173; DB 2; Length 409;
Best Local Similarity	48.0%;	Pred. No. 8.1e-79;
Matches 223; Conservative	63;	Mismatches 111; Indels 68; Gaps 8;
QY 23	GDICDPNPGENGICL-----PGLAVGSFSCCECPDGFDPNCSSVVEVASDEEPTSA 75	
DB 3	GDICDSSLCLNGGTCLDDDPQKP-----PFLCPEGFTGLCNE-----TEK 45	
QY 76	GCPETPNCHNGTCE-ISAAYRGDTFTGYVCKCPKGFNIRHCHNINECEVEPKNGGIC 134	
DB 46	GCPETPNCHNDACEVIDDAHRGDEVTEYICPCPHGYTIIH----- 86	
QY 135	TDLVAVNSCECPPEFMRNOCYKSGPLGIEGGIISNOOITASSSTHRAFLGLOKWPYYA 194	
DB 87	-----CEIICNAPLDMEFGALADPDISSSMHLGIMGIORAPELA 127	
QY 195	RLNKGLINAWTAENDRRKRWIQLMQRMRVYGVITOGARRIGSPYIKFYKIAYSND 254	
DB 128	RLNRAGIVANWMTSNYDR-NPMIQLVNLRRMRVGVVTOGASRAGSAEYMKTFKVAYSTD 186	
QY 255	GKTYWAKYKKGWNE--DMYFRGINDNNTYANSFTPIRAQVYRLYPOVCRRHGCTLRMEL 312	
DB 187	GRFRQF--TQGAESGDKITFMGLNDSGLKVLFEVPLEVQVYRLVPIIICHGCTLRREL 244	
QY 313	LGCELGGCSEPLGKSGHIQDOYQITASSIFRTLNDMFTWERPKARLDKOGKVNAMTSGH 372	
DB 245	LGCELGGCAEPDLKQNTIPNKQITASSFRYTWGLAFSMFYFARLDQGFNMTAS 304	
QY 373	NDOSQWLOVDLVPVYVTGILITOGADFGHVOFVGSYKILAYSNDGEHMTVYODEKORKDK 432	

Db 305 NSASEMLQIDLQSGQRVVTGIIITOGARDFGHIOYVAARKVAYSDGVSWTEYRDQALGEK 364
 QY 433 VFOGPNFDNTHRRKVVIDPPIYARHRIILPMSWGRIITLASELLGC 477
 Db 365 IFPGNLDNNSHKRMETPEPLTFVRLIPVAMHNRITLRELLGC 409

RESULT 2

A36479
 milk fat globule membrane protein - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A36479
 R:Stubs, J.D.; Lekutis, C.; Singer, K.L.; Bul, A.; Yuzuki, D.; Srinivasan, U.; Parry, G
 Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990
 A>Title: CDNA cloning of a mouse mammary epithelial cell surface protein reveals the ex
 A:Reference number: A36479; MUID:91046008
 A:Accession: A36479
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-463 <STU>
 A:Cross-references: GB:M38337; NID:9199142; PIDN:AAA39534.1; PID:9199143
 C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
 C:Keywords: membrane protein
 F:28-60/Domain: EGF homology <EG1>
 F:68-107/Domain: EGF homology <EG2>
 F:147-303/Domain: discoidin I amino-terminal homology <DN1>
 F:307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match 43.2%; Score 1156; DB 1; Length 463;
 Best Local Similarity 48.5%; Pred. No. 1.7e-77;

Matches 222; Conservative 72; Mismatches 142; Indels 22; Gaps 8;

QY 23 GDICDPNCEGICLPGLAVGSCPCPDGFTDPNCSSVVEASDEEPTSAQCTPNP 82
 Db 25 GPCDSSLCLNGGTCLTG-QDNDIYCLCPGFTGLVNE-----TERGCSCTNP 72
 QY 83 CHNGTCSEIS-EAYRGDTFTGYVCKPCRGFNGIHCOHINNECEVEPCCKNGICTDLVANY 141
 Db 73 CINDAKCLVTLDTRQDIFTEYICCPVSGIHCEETETNYNDL---GEYFTTAVPNT 129
 QY 142 SCECPGEF--MGRNCOYKSGPLGEGIIISNOQITASTHRLALGLOKWPYVRLKK 199
 Db 130 AVTPAPPTDLSNNLASRSTQLGMEGAIADSOISASIVYGFGLDQWGPDLRLKLT 189
 QY 200 GLINAVTAENDRKRWIOINLQRMKRVTVITGAKRIGSEYIKFKIAYSDNGKTA 259
 Db 190 GIYVAMHASNYDS-KPWIOVNLRRKRVSGVMTQASRAGRAEYLTKEFVAYSIDGRKFE 248
 QY 260 MKVVGTDNDVFRGINDNNTPIANSFTPIKAQVRLYPQVCRHRTLRMLLCELSG 319
 Db 249 FIQDE-SGGDKFELNLDNNSLTKVMFNPTLEAOYIRLYPVSCHGCTLRRELLCEELHG 307
 QY 320 CSEPLGMSGHIQDYOTASSIFRTLNDMFTWEPKARLDKQGVNMTVSGHNDQSOVL 379
 Db 308 CLEPLGLKNNITIPDSQMSASSSYKTWNLRAFQWYHLGRLDNQGKINMTVQSNKAKML 367
 QY 380 QVDLVPKRVGIIITGAKDFGHVQVGSYKLAYSNDGEHMTVYODEKQRKDVQGNFD 439
 Db 368 QVDLGTORVGIITGQARDFGHIQYVESYKVAHSDQVQWTVY--EEQSSKVFQGMLD 425
 QY 440 NDTHRKNVIDPPIYARHRIILPMSWGRIITLASELLGC 477
 Db 426 NNSHKNIPEKPFMARIVRLPVSWMHNRITLRELLGC 463

RESULT 3

JC4915
 ags protein precursor - rat
 N:Alternate names: O-acetyl-Gd3 ganglioside
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 20-Jun-2000

C:Accession: JC4915
 R:Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanae, Y.
 Biochem. Biophys. Res. Commun. 225, 932-938, 1996
 A>Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.
 A:Reference number: JC4915; MUID:96374422
 A:Accession: JC4915

A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-427 <OGU>
 A:Cross-references: DDBJ:D84068; NID:91620006; PIDN:BA12210.1; PID:91620007
 A:Experimental source: CST cell
 C:Comment: This protein is required for the O-acetylation of distalganglioside stial
 C:Genetics:
 A:Gene: ags

C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF hom
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:28-60/Domain: EGF homology <EG1>
 F:68-107/Domain: EGF homology <EG2>
 F:110-267/Domain: discoidin I amino-terminal homology <DN1>
 F:271-427/Domain: discoidin I amino-terminal homology <DN2>

Query Match 43.1%; Score 1154; DB 2; Length 427;
 Best Local Similarity 48.5%; Pred. No. 2.1e-77;

Matches 221; Conservative 62; Mismatches 119; Indels 54; Gaps 7;

QY 23 GDICDPNCEGICLPGLAVGSCPCPDGFTDPNCSSVVEASDEEPTSAQCTPNP 82
 Db 25 GPCDSSLCLNGGTCLTG-QDNDIYCLCPGFTGLVNE-----TERGCSCTNP 72
 QY 83 CHNGTCSEIS-EAYRGDTFTGYVCKPCRGFNGIHCOHINNECEVEPCCKNGICTDLVANY 141
 Db 73 CFHAKCLVTBDTRQDIFTEYICCPVSGIHCE----- 108
 QY 142 SCECPGEF--MGRNCOYKSGPLGEGIIISNOQITASTHRLALGLOKWPYVRLKK 201
 Db 109 -----LG-----CSTKLGEGAIADSOISASSVYGFGLDQWGPDLRLKLTGI 154
 QY 202 INAVTAENDRKRWIOINLQRMKRVTVITGAKRIGSEYIKFKIAYSDNGKTA 261
 Db 155 VNMTASTSYDS-KWIOVNLRRKRVSGVMTQASRAGRAEYLTKEFVAYSIDGRREFT 213
 QY 262 KVKCTNEDVFRGINDNNTPIANSFTPIKAQVRLYPQVCRHRTLRMLLCELSGS 321
 Db 214 QDESGTDKEMGNODNNSLTKVMFNPTLEAOYIRLYPVSCHGCTLRRELLCEELHGS 273
 QY 322 EPLGMSGHIQDYOTASSIFRTLNDMFTWEPKARLDKQGVNMTVSGHNDQSOVL 381
 Db 274 EPLGLKNNITIPDSQMSASSSYKTWNLRAFQWYHLGRLDNQGKINMTVQSNKAKML 333
 QY 382 DLAVPKRVGIIITGAKDFGHVQVGSYKLAYSNDGEHMTVYODEKQRKDVQGNFD 441
 Db 334 DLGTQKRVTVITGQARDFGHIQYVESYKVAHSDQVQWTVY--EEQSSKVFQGMLD 391
 QY 442 THKKNVIDPPIYARHRIILPMSWGRIITLASELLGC 477
 Db 392 SHKKNIFEKPFMARIVRLPVSWMHNRITLRELLGC 427

RESULT 4

S74211
 PAS-6/7 protein precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
 C:Accession: S74211; S78114; S24181
 R:Harregard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.
 Eur. J. Biochem. 240, 628-636, 1996
 A>Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat g
 A:Reference number: S74211; MUID:97008954
 A:Accession: S74211
 A:Molecule type: mRNA
 A:Residues: 1-427 <HVA>
 A:Cross-references: EMBL:X91895; NID:91632778; PIDN:CAA62997.1; PID:91632779

F;2193-2345/Disulfide bonds: #status predicted

Query Match	24.38;	Score 649;	DB 1;	Length 2351;
Best Local Similarity	44.28;	Pred. NO. 1.8e-39;		
Matches 144;	Conservative 46;	Mismatches 120;	Indels 16;	Gaps 5;

[illegible]

RESULT 11
A47285 milk fat globule protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change 28-May-1999
C:Accession: A47285
R:Laococca, D.; Peterson, J.A.; Urrea, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
Cancer Res. 51, 4994-4998, 1991
A:Title: A-M-r 46,000 human milk fat globule protein that is highly expressed in human b
A:Reference number: A47285; MUID:91371351
A:Accession: A47285
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <IAR>
A:Cross-references: CB:SS6151; NID:g235396; PIDD:AA019771.1; PID:g235397
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homology
P1-56/Domain: discoidin I amino-terminal homology (fragment) <DNI>
P:60-218/Domain: discoidin I amino-terminal homology <DN2>

Query Match	24.0%;	Score 643;	DB 2;	Length 218;
Best Local Similarity	56.8%;	Pred. NO. 3.2e-40;		
Matches 117;	Conservative 30;	Mismatches 59;	Indels 0;	Gaps 0;

[illegible]

RESULT	12
T42763	

C:coagulation factor VIII precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C:Accession: T42763
R:Lollar, P.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z22269
A:Accession: T42763
A:Status: preliminary; translated from GR/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-2133 <L>
A:Cross-references: EMBL:U49517; NID:g1511633; PID:g1511634; PIDN:AA06705.1
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A;
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
F:23-349/Domain: ferroxidase repeat homology <FOX1>
F:402-750/Domain: ferroxidase repeat homology <FOX2>
F:1458-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match	23.9%;	Score 639;	DB 2;	Length 2133;
Best Local Similarity	40.0%;	Pred. No. 8.8e-39;		
Matches 144;	Conservative 52;	Mismatches 136;	Indels 28;	Gaps 6

OY	122	ECXEYPCNKSGICGDIUAVANS	SCCCPGEFMRGNQYKSCSP	IEGIIISNOITASSPHR	181	
Db	1800	ECJLGEHLOQMSTTELV	YSKE-----	CAQPLGMSAGRDRDQITASQY	1844	
OY	182	ALFGLQWYPIYVARLANKGL	LINMTAENDRMKRWIOJLQ	KRRVYGITOGAKRIGSP	241	
Db	1845	-----GOWAPRLAHL	SSGSIINAMSTRDP----	HSWIKDILLAPMILHIMTQ	AROKRFS 1896	
OY	242	EYIKFYKIAVSNCKCTAAM	KVKGTEDEMYFRGNIDNN	TPYANSEFTPIKQVRLPYOV	301	
Db	1897	LYISOFIIMSJLORNMQSI	RGNSTGTGLVFFGNDV	ASIKINIINPIYVARIRLHP	TH 1956	
OY	302	CRRHCTLMELLGCELSG	CESEPLGKMSGHIQYQIT	ASSIEFTLMDMFWEPKARL	DLK 361	
Db	1957	YSIRSTLRMELMGDILN	SCSMPGLMONKAISDSQ	ITASS--HLSNIFATSP	SOARHL 2013	
OY	362	QGRNANMNSGHNDSQ	LOJYDLVLPKRYGII	TQOAGKDEGHQYFAGSK	LAYSNDGEHMT 421	
Db	2014	QGRNANMRPRKSSAEHL	QYDLQKTYKVGITITQ	GYKSLSSMYKELVLVSS	SODGRWMT 2073	
OY	422	VYUDEKORKDKVFOG	NDFNDTHRRKNV	IDPPYVARHRIILPMSW	GRITLASSELLGTEEB 481	
Db	2074	LELDQGH--IKVEFG	NDSDSTPRVANAIDP	LETPTRYLRHPTSM	AOHILRLREVLEGC	ENAD 2131

RESULT 13
A44258
Factor VII-associated gene B hypothetical protein - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_rev:10 30-Sep-1993 #text_change 18-Jun-1999
C:Accession: A44258
R:Levinson, B.; Kenwright, S.; Gamel, P.; Fisher, K.; Gitschier, J.
Genomics 14, 585-589, 1992
A:Title: Evidence for a third transcript from the human factor VIII gene.
A:Reference number: A44258; PMID:93052386

Query Match 17.48; Score 464.5; DB 2; Length 216;

Best Local Similarity 46.7%; Pred. No. 4.1e-27;
Matches 99; Conservative 31; Mismatches 75; Indels 7; Gaps 3;

```
QY 271 VERNDINNTFYANSFTPPKAOVRLYPOVCRHCHTLRMELLGCELSGSGSEPLGKSGH 330
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 VEFNVDSGGIKHNEFPPIIARIIRLPHYISIRSTLRMLMGCCDLSGSMPLGMSKA 68
QY 331 IQDYQITASSIFRTLNMDMF-TWEPRKARLDKQGVNAWTSNGHDSQWMLQVLDLVPYKV 389
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 ISDAQITASSYF-----TMMFATNSPKARLHLQGRSNAMRPQVNNKPEWLOVDFQKTKMY 124
QY 390 TGIITGAKDFGVQVQVSGYKLAYSNDGEHWTVYQDEKQKRDYFQGNFNDTHRRKNVID 449
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 TGVTGQGVKSLTSMYKEFLISSQDGHQWTLF--FQNGKVKVFGQGNDSFTPVVNSLD 182
QY 450 PPIYARHRLPMSWYGRITLASLLGCTEEB 481
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 183 PPLLRTRLRHPSWVHOIALRMEVLGCEAD 214
```

RESULT 14

J00948

A5 antigen precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JH0466; J00948
R:Takeagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
Neuron 7, 295-307, 1991
A:Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homology
A:Reference number: JH0466; MUID:91337438
A:Accession: JH0466
A:Molecule type: mRNA
A:Residues: 1-927 <TAK>
A:Cross-references: GB:D10467; GB:D01077; NID:g222962; PIDN:BAA01260.1; PID:g222963
A:Experimental source: tadpole, brain
A:Note: this protein has motifs homologous to complement components C1r and C1s and to C
C:Comment: This protein is a neuronal cell surface molecule involved in the neuronal rec
C:Superfamily: Xenopus A5 antigen; C1r/C1s repeat homology; discoidin I amino-terminal R
C:Keywords: duplication; glycoprotein; transmembrane protein
F:1-21/Domain: signal sequence predicted <SIG>
F:22-927/Product: A5 antigen #status predicted <ASA>
F:27-118/Domain: C1r/C1s repeat homology <C1R1>
F:147-262/Domain: C1r/C1s repeat homology <C1R2>
F:214-424/Domain: discoidin I amino-terminal homology <DN1>
F:430-584/Domain: discoidin I amino-terminal homology <DN2>
F:646-812/Domain: MAM homology <MAM>
F:861-883/Domain: transmembrane #status predicted <TM>
F:150,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.8%; Score 450; DB 1; Length 927;

Best Local Similarity 31.2%; Pred. No. 2.5e-25;
Matches 155; Conservative 58; Mismatches 214; Indels 70; Gaps 23;

```
QY 12 GLSLGVQFGGDIQDPNCPENGICLPGLAVSF--SCPCPDGFTDPNCSVY-EVASD 68
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 133 GFSIRYEFKFGPCSRNFTSSNGVIKSPRYEKYPNALCTYIIIPAKMQEIVLEFESR 192
QY 69 EEPRTSAGPCTPNCHNGTCE-----ISEAYRG-DTFIGYVC--KCP---RGNGI--- 114
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 193 ELEADSNAP-----GGQICRYWLGIMWGFQVGHIGRYGQNTPRGVRSTGLISM 245
QY 115 --HCQHNINECEVPCNKGICITDLVANYSCECPGEFMRNC--QYKSGPLIEGGIIS 170
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 IFHDSAI-----AKEG-----FFANFS-----VVGQNTDEDFQCKEALGSEGEIH 287
QY 171 NQQTASTHRALEGLQKVPYVYARLKKGLINAWTAENDRKRMVQIQLQKRMKRTGV 230
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 288 FDQISVSQSYF---MMSAERSRLNY--VENGWTPGE--DTVKEWLOVDLEMLRFSGI 339
QY 231 ITQGA--KRISPEYIKFYKLAYSNDGKTWAMYKVGKTNDVFRGINDNTFYANSFTP 288
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 340 GTQGAISKETKKKKFYKYSKYVDISSNGEDM--ITLKQGNKHLVFTGNTDATDVYVRFK 397
```

```
QY 289 PIKAOVRLYPOVCRHCHTLRMELLGCELSG--GSEPLGMSGHIODYQTASSIFRTLN 346
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 398 PYITFRVRLRPYTWENGISLRFELIGCKITTPYCSRMGLAMSGLISDOITASS-----Q 452
QY 347 MDMEFWEPKARL--DKQGVNAWTSNGHDSQWMLQVLDLVPYKVYITGQANDFGHVQ 404
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 453 VDR-NWVPELARLVTSRSGMALPPSNTHPYTKEMWLOIDLAEEKIVRGVYIGGKHKENKV 511
QY 405 FVGSYKLAYSNDGEHWTVYQDEKQKRDYFQGNFNDTHRRKNVIDPPIYARIRILP-W 463
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 512 FMKFRKIGYSNNGTEWEMTMDSSKNPKTEEGNTVYDPELRTF-AHTTGFIRILPERA 570
QY 464 WYGRITLASLLGCTEEB 480
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 571 SASGLALRRELLGCEVE 587
```

RESULT 15

S52093

hemocytin - silkworm

N:Alternate names: humoral lectin

C:Species: Bombyx mori (Silkworm)

C:Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 17-Mar-1999

C:Accession: S52093; S70920

R:Kotani, E.; Yamakawa, M.; Iwamoto, S.; Tashiro, M.; Mori, H.; Sumida, M.; Matsubara

Biochim. Biophys. Acta 1260, 245-258, 1995

A:Title: Cloning and expression of the gene of hemocytin, an insect humoral lectin wh

A:Reference number: S52093; MUID:95178544

A:Accession: S52093

A:Molecule type: mRNA

A:Residues: 1-313 <KOT>

A:Cross-references: EMBL:D29738

R:Mori, H.

submitted to the EMBL Data Library, April 1994

A:Reference number: S70920

A:Accession: S70920

A:Molecule type: mRNA

A:Residues: 1-1566, 'S', 1568-3133 <MOR>

A:Cross-references: EMBL:D29738; NID:g474967; PID:d1006722; PID:g664884

Query Match 14.0%; Score 374; DB 2; Length 3133;

Best Local Similarity 24.7%; Pred. No. 3.9e-19;
Matches 148; Conservative 70; Mismatches 174; Indels 206; Gaps 31;

```
QY 32 ENGIGD-----LPGIAGVSFSCQPDGFTDPNCSVYEVASD-BEPTSAAGPCTPNCH 84
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 715 ENGLVCEEPKTCPPDEIKVY-CECEE--PDSTSPVTVTSSEASSEVSTTLATTT--- 767
QY 85 NGGTCEISEAYRGDPIFIGVCKCPRGFNGIHCOH-----NINECEVPCNKGICITD-L 137
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 768 --SRCPGEGVYQACV-----KCDR-----LCDHKKTLIAKGRISLSEYCDG--CYDES 813
QY 138 VANYSC-----C-----PGEFGRNCOYKC----- 158
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 814 VASNGCEGSSRMDEKTCVPRVDCIYNDQIYKPGVYESGC-IKCQCLDNLSCDSKD 872
QY 159 ----- 166
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 873 CVSLNIPHOGSHLPYIYRPVSTTTTSTTTTSTTTTSTTTTPTPEPTETTTETVYPLIKS 932
QY 167 GI-----ISNQQTASTHRALEGLQKVPYVYARLKKGLIN--- 203
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 933 TVSPPECSPDVYIDLVMGDPLPPTARSASSESEIFA-----PHNRLN-RGPTNSGA 986
QY 204 -AWTAENDRKRMVQIQLQKRMKRTGVITGQAKKIGSP--EYIKFYKLAYSNDGKTWA 259
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 987 GSMNPRVND-KQYLOVELPREPIYGVLO-----GSPIDQVYTSYEIYWGDDGNTFS 1040
QY 260 MKYKVTN-EDWVFGNDINNTFYANSFTPPKAOVRLYPOVCRHCHTLRMELLGCG--- 315
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1041 --TVGDPGDKPKIEFGPIDNTHPVKQMISPIEAKVVRIRPLTHDEISLRLEITIGCAEP 1098
```

Wed May 23 06:29:19 2001

us-09-237-981-14_copy_33_513.rpr

Page 9

[illegible]

Search completed: May 23, 2001, 06:12:43
Job time: 77 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:18:52 ; Search time 34.32 Seconds
(without alignments)
480.096 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513

Perfect score: 2675
Sequence: 1 MKRSYAVMLVGLSLGVPOF.....MSWYGRITLASELLGCTEE 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1173	43.9	409	1	MEGM_PIG
2	1156	43.2	463	1	MEGM_MOUSE
3	1154	43.1	427	1	MEGM_RAT
4	1143	42.7	427	1	MEGM_BOVIN
5	1020.5	38.1	387	1	MEGM_HUMAN
6	675.5	25.3	2224	1	FA5_HUMAN
7	670	25.0	2319	1	FA8_MOUSE
8	657.5	24.6	2211	1	FA5_BOVIN
9	649	24.3	2351	1	FA8_HUMAN
10	639	23.9	2133	1	FA8_PIG
11	465.5	17.4	923	1	NRPL_HUMAN
12	461.5	17.3	914	1	NRPL_CHICK
13	458.5	17.1	923	1	NRPL_MOUSE
14	457	17.1	925	1	NRPL_RAT
15	457	17.1	931	1	NRPL_MOUSE
16	453	16.9	931	1	NRPL_HUMAN
17	450	16.8	928	1	NRPL_XENLA
18	447.5	16.7	922	1	NRPL_RAT
19	374	14.0	3133	1	HMCT_BOVINO
20	326	12.2	1064	1	FBP1_STRPU
21	313.5	11.7	2524	1	NOTC_XENLA
22	312	11.6	2531	1	NTC1_RAT
23	311.5	11.5	2531	1	NTC1_MOUSE
24	307	11.5	2437	1	NOTC_BRARE
25	303.5	11.3	570	1	FBP3_STRPU
26	299.5	11.2	2703	1	NOTC_DROME
27	295.5	11.0	280	1	XLRL_FUGRU
28	295	11.0	2444	1	NTC1_HUMAN
29	292.5	10.9	1408	1	SERR_DROME
30	291.5	10.9	2318	1	NTC3_MOUSE
31	287.5	10.7	589	1	DL13_RAT
32	287	10.6	723	1	DL11_HUMAN
33	282.5	10.6	592	1	DL13_MOUSE

34	280	10.5	224	1	XLRL_MOUSE
35	275	10.3	714	1	DL11_RAT
36	274.5	10.3	833	1	DL_DROME
37	271.5	10.1	618	1	DL13_HUMAN
38	268	10.0	473	1	FP2_MITGA
39	267	10.0	473	1	XLRL_HUMAN
40	267	10.0	722	1	DL11_MOUSE
41	265	9.9	4590	1	FATH_HUMAN
42	262.5	9.8	2139	1	GRB_DROME
43	259.5	9.7	1480	1	SLIT_DROME
44	259	9.7	1964	1	NTC4_MOUSE
45	245	9.2	383	1	DLK_HUMAN

ALIGNMENTS

RESULT	ID	MEGM_PIG	STANDARD	PRT	409 AA
MEGM_PIG	1	MEGM_PIG			
AC	P79385				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	15-JUL-1998	(Rel. 36, Last sequence update)			
DT	30-MAY-2000	(Rel. 39, Last annotation update)			
DE	LACTADHERIN (MILK FAT GLOBULE-EGF FACTOR 8) (MEG-8) (MEGM) (SPERM				
DE	SURFACE PROTEIN SP47) (PP47).				
GN	MEG8.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.				
ON	NCBI_TaxID=9823;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Testis;				
RA	Ensslin M.A.;				
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.				
CC	- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIP BINDING. ZONA PELLUCIDA-				
CC	- BINDING PROTEIN.				
CC	- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.				
CC	- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND				
CC	SPERMATOZOAN.				
CC	- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.				
CC	- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.				
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	EMBL: Y11683; CAAT2379.1; -				
DR	HSP: P00740; IIXA.				
DR	InterPro: IPR000421; -				
DR	InterPro: IPR000561; -				
DR	Pfam: PF00008; EGF_2.				
DR	Pfam: PF00754; F5_F8_Type_C; 2.				
DR	PROSITE: PS00022; EGF_1; 2.				
DR	PROSITE: PS01186; EGF_2; 2.				
DR	PROSITE: PS01285; FAS8C_1; 2.				
DR	PROSITE: PS01286; FAS8C_2; 2.				
KW	Glycoprotein; Repeat; EGF-like domain.				
FT	DOMAIN 2				
FT	DOMAIN 44				
FT	DOMAIN 88				
FT	DOMAIN 91				
FT	DOMAIN 252				
FT	DOMAIN 409				
FT	SITE 67				
FT	SITE 69				
FT	DISULFID 6				
FT	DISULFID 11				
FT	DISULFID 17				
FT	DISULFID 29				
FT	DISULFID 31				
FT	DISULFID 40				
FT	DISULFID 91				
FT	DISULFID 247				
FT	DISULFID 234				
FT	DISULFID 238				

FT DISULFID 252 409 BY SIMILARITY.
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 409 AA; 45725 MW; BOC07AF80029927A CRC64;

Query Match 43.9%; Score 1173; DB 1; Length 409;
 Best Local Similarity 48.0%; Pred. No. 8.2e-82;
 Matches 223; Conservative 63; Mismatches 111; Indels 68; Gaps 8;

OY 23 GDICDPNPNCEGICL-----PGLAVGSFSCPCPDGFTDPCSSVVEASDEEPTSA 75
 DB 3 GDFCDSSICLNGICLIDDPQPKP-----FHICGPEGTGLICNE-----TER 45
 OY 76 GCPTRPNCHNGTCE-ISEAYREDFTIGYVCKCPGFNGICHOHNINECEVEPCNKGIC 134
 DB 46 GPCFPNCPCHNDACEVIDDAHGRDVFTEYICKCPHGYTDIH----- 86
 OY 135 TDLVANSCCECPGFEGMRCNOYKCGPLGIEGIIISNOITFASSTHRAFLFGLOKWPYYA 194
 DB 87 -----CEIICNAPLGHEGICALDFOISASSMHLGFMGLQRRAPELA 127
 OY 195 RLNRKGLINAMTAEANDRMKRWIQLNQRKRVYITOGARRISPEYIKFYKIVASND 254
 DB 128 RLHAGIYNAMWTASNDR-NPWIOVNLRRMRVTVGTGASRAGSAEYMKTFKVAYSTD 186
 OY 255 GKTNAMKVKGTNE--DMVFRGNIDNNTPYANSEFPPIKAQVYRLYPQCRHCHTBMEL 312
 DB 187 GRKQF---IQAEBSGSKITFMGNDNSGLKVNLEFVPELVQYRVLPIICHRGCTLREL 244
 OY 313 LGCELSCGSEPLGKSGHIDYQITASSIFRLINMDFTWEPKRAKLOKGVNMTSGH 372
 DB 245 LGCEISGCAEPLGLKMDTIPNKQITASSFYRTGLSAFSMYPFARLNOGCFNMTAQS 304
 OY 373 NDOSOMLOVDLVPFKVTGIIITOGADFGHVQVSGKATYANDEHMTVODEQRQDK 432
 DB 305 NSASEMLQIDIGSORRYTGIIITOGADFGHIQVAAKYAISDDGYSTETRDQALEGK 364
 OY 433 VEOGNFMDTHRKNVDPPIYARHIRLLPMSWYGRITLASELGC 477
 DB 365 IFPGMLDRNHSKMKMFETPLFTRFVRLPYAMHNRITLREVELLGC 409

RESULT 2
 MFGM_MOUSE STANDARD; PRT; 463 AA.
 AC P21956; P97800;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (MFGM)
 DE (SPERM SURFACE PROTEIN SP47) (MP47).
 GN MFG8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 11)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-35.
 RC TISSUE-Mammary gland;
 RX MEDLINE=91046008; PubMed=2122462;
 RA Stubbs J.D., Lekutis C., Singer K.L., Bui A., Yuzuki D.,
 RA Sriyasan U., Parry G.;
 RT "cDNA cloning of a mouse mammary epithelial cell surface protein
 RT reveals the existence of epidermal growth factor-like domains linked
 RT to factor VIII-like sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
 RN 12)
 RP SEQUENCE OF 23-463 FROM N.A.
 RC TISSUE-Testis;
 RA Engslin M.A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN PROSPHOLIPID BINDING. ZONA PELLUCIDA-

CC BINDING PROTEIN.
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND
 CC SPERMATOZOAN.
 CC -1- DEVELOPMENTAL STAGE: MRNA EXPRESSION IS DETECTABLE IN MAMMARY
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M38337; AAA39534.1; -;
 DR EMBL: Y11684; CAAT2380.1; -;
 DR PIR: A36479; A36479.
 DR HSP: P00740; 11XA.
 DR MGD: MGI:102768; M1ge8.
 DR InterPro: IPR000421; -;
 DR InterPro: IPR000561; -;
 DR InterPro: IPR001438; -;
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00754; F5_FB_Type_C_2.
 DR PRINTS: PR00010; EGF_HOOD.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 DR Signal: Glycoprotein; Repeat; EGF-like domain; Milk.
 FT CHAIN 1 22
 FT SIGNAL 1 22
 FT CHAIN 23 463
 FT DOMAIN 24 61
 FT DOMAIN 64 108
 FT DOMAIN 148 303
 FT DOMAIN 308 463
 FT SITE 87 89
 FT DISULFID 28 39
 FT DISULFID 33 49
 FT DISULFID 51 60
 FT DISULFID 68 79
 FT DISULFID 73 96
 FT DISULFID 98 107
 FT DISULFID 148 303
 FT DISULFID 290 294
 FT DISULFID 308 463
 FT CARBOHYD 61 61
 FT CARBOHYD 266 266
 FT CARBOHYD 316 316
 FT CARBOHYD 426 426
 FT CONFLICT 30 30
 FT CONFLICT 35 35
 FT CONFLICT 110 147
 FT CONFLICT 168 168
 FT CONFLICT 196 196
 FT CONFLICT 309 309
 FT CONFLICT 395 395
 SO SEQUENCE 463 AA; 51465 MW; D78B6C6FBBA724D CRC64;

Query Match 43.2%; Score 1156; DB 1; Length 463;
 Best Local Similarity 48.5%; Pred. No. 1.8e-80;
 Matches 222; Conservative 72; Mismatches 142; Indels 22; Gaps 8;
 OY 23 GDICDPNPNCEGICLPLAVGSFSCPCPDGFTDPCSSVVEASDEEPTSA 82
 DB 25 GDFCDSSICLNGICLIDDPQPKP-----FHICGPEGTGLICNE-----TER 72
 OY 83 CHNGCTCEIS-EAYRGDTFTIGYVCKCPGFNGIHCCHOHNINECEVEPCNKGICITDLVANY 141

```

DB 73 CYNDACKLVTLDTORGDIFTEYICCPVGYSGHCEETETNYNND--GEYFTTAVPNT 129
QY 142 SCSCPEEF--MGRCQYKCSGPLGEGGITSNOQTASTSTRALFGLQKWPYARLKK 199
DB 130 AVPTPAPTPDLNNLASRCSSTOLGMEGAIADSOISASVYWGFGLORMPELARLYRT 189
QY 200 GLINAMTAENDMRKMIQINLQKRMVTVGTITGAKRIGSPETIKFYKIAVSDGKTWA 259
DB 190 GIYANMAHASYDS-KPMIOVNLKRMVSGVMTGASRAGRAYLKTFRKVAVSLDGRKFE 248
QY 260 MYKVTGNTEDMVRGNIDNNTPRANSFTPRKAOYVRLYPOVCRHCTLRMEILGCELSG 319
DB 249 FIDF--SGGKREFGLNIDNNSLKVNNFNPTLPAQYIRLIPVCHGCTLRPELLGCELHG 307
QY 320 CSPELGKSGHIDYOITASSIFRTLMDMFTWEPKARLDKQGVNMTSGHNDQSGML 379
DB 308 CLEPLGLKNNITPDSQMSASSYKTNLRAFQWYPLGLRDLNQGKINAMTAQSNAGKEML 367
QY 380 QVDDLVPKVTGIIITGAKDFGVGVGSYKLAYSMDGSHWTVYODEKORKKXVQGNFMD 439
DB 368 QVDLGTORQVYTGIIITGAKDFGHIOYVESYKVAHSDGVQWTVY--EEQSSKVFQGNLD 425
QY 440 NDRHRKNVIDPPIYARHIRILPMSWYGRITLASELLGC 477
DB 426 NSHKKNIFEKPMARYVRLPLVSMNRITLRLLELLGC 463

```

RESULT 3
MEGM_RAT STANDARD: PRT: 427 AA.

```

AC P70490;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-EB) (O-
DE ACETYL G33 GANGLIOSIDE SYNTHASE) (AGS) (MEGM).
GN MFGEB OR AGS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=96374422; Pubmed=8780713;
RX Ogura K., Nara K., Watanabe Y., Kohno K., Tai T., Sanai Y.;
RT "Cloning and expression of cDNA for O-acetylation of G33
RT ganglioside."
RT Biochem. Biophys. Res. Commun. 225:932-938(1996).
RL -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. SEEMS TO
PARTICIPATE IN THE O-ACETYLTATION OF G33 GANGLIOSIDE SIALIC ACID.
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: SPLEEN, LUNG, HEART, BRAIN AND MUSCLE.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D84068; BAAL2210.1; -
CC HSSP: P00740; IIXA.
CC InterPro: IPR000421; -
CC InterPro: IPR000561; -
CC InterPro: IPR001438; -
CC Pfam: PF00008; EGF_2.
CC Pfam: PF00754; F5_F8_type_C_2.
CC PRINTS: PR00010; EGFBL00D.

```

```

DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FAS8C_1; 2.
DR PROSITE: PS01286; FAS8C_2; 2.
KW Signal; Glycoprotein; Repeat; EGF-like domain; Milk.
FT CHAIN 1 23 427
FT DOMAIN 24 61
FT DOMAIN 64 108
FT DOMAIN 111 267
FT DOMAIN 272 427
FT DISULFID 28 39
FT DISULFID 33 49
FT DISULFID 51 60
FT DISULFID 68 79
FT DISULFID 73 96
FT DISULFID 98 107
FT DISULFID 111 267
FT DISULFID 254 258
FT DISULFID 272 427
FT SITE 87 89
FT CARBOHYD 61 61
FT CARBOHYD 230 230
FT CARBOHYD 280 280
FT CARBOHYD 390 390
SQ SEQUENCE 427 AA; 47413 MW; EA8C8631F3EE6047 CRC64;

```

Query Match 43.1%; Score 1154; DB 1; Length 427;
Best Local Similarity 48.5%; Pred. No. 2,4e-80;
Matches 221; Conservative 62; Mismatches 119; Indels 54; Gaps 7;

```

QY 23 GDICDPNCPENGICLPLGLAVSFSCEPDGDTFPCSSVVEVASDEEPTAGCTPMP 82
DB 25 GDFCDSSSLCNGTGCLMG-QDNNIYLCPEGFGVLCNE-----TEKGCSPNP 72
QY 83 CHNGCTEISE-AVRGCTFTGYCKCPKRGNGIHQCHNINECEVCEKNGGICDLPVNY 141
DB 73 CFHDAKCLVETDQRGDIFTEYICCPVGYSGIHCE----- 108
QY 142 SCCEPGEFMGRCQYKCSGPLGEGGITSNOQTASTSTRALFGLQKWPYARLKKGL 201
DB 109 -----LG-----CSTKLGLGEGAIADSOISASVYWGFGLORMPELARLYRTGI 154
QY 202 INAMTAENDMRKMIQINLQKRMVTVGTITGAKRIGSPETIKFYKIAVSDGKTWAMY 261
DB 155 VNAMTASYSYDS-KPMIOVDFLRKMVSGVMTGASRAGRAYLKTFRKVAVSLDGRFEEI 213
QY 262 KYKGTNEDMVRGNIDNNTPRANSFTPRKAOYVRLYPOVCRHCTLRMEILGCELSGS 321
DB 214 QDESCTGDKERMGQNDNNSLKNINFNPTLEQYIRLIPVCHGCTLRPELLGCELHGS 273
QY 322 EPLGKSGHIDYOITASSIFRTLMDMFTWEPKARLDKQGVNMTSGHNDQSGMLQV 381
DB 274 EPLGLKNNITPDSQMSASSYKTNLRAFQWYPLGLRDLNQGKINAMTAQSNAGKEMLV 333
QY 382 DLVPTKVTGIIITGAKDFGVGVGSYKLAYSMDGSHWTVYODEKORKKXVQGNFMD 441
DB 334 DLGTOKRYVTGIIITGAKDFGHIOYVASKVAHSDGVQWTVY--EEQISKVFQGNLDNN 391
QY 442 THRKNVIDPPIYARHIRILPMSWYGRITLASELLGC 477
DB 392 SHKKNIFEKPMARYVRLPLVSMNRITLRLLELLGC 427

```

RESULT 4
MEGM_BOVIN STANDARD: PRT: 427 AA.

```

AC Q95114; Q27959; P79344;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-EB)

```


DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (HMF8)
 DE (BREAST EPITHELIAL ANTIGEN BA46) (MFGM) [CONTAINS: MEDIN].
 GN MFG8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Breast, and Breast carcinoma;
 RX MEDLINE=96213908; PubMed=8639264;
 RA Couto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Peterson J.A.;
 RT "Cloning and sequence analysis of human breast epithelial antigen
 RT BA46 reveals an RGD cell adhesion sequence presented on an epidermal
 RT growth factor-like domain.";
 RL DNA Cell Biol. 15:281-286(1996).
 RN [2]
 RP SEQUENCE OF 170-387 FROM N.A.
 RC TISSUE-Mammary gland;
 RX MEDLINE=91371351; PubMed=1909932;
 RA Larocca D., Peterson J.A., Urrea R., Kuniyoshi J., Bistrain A.M.,
 RA Ceriani R.L.;
 RT "A Mr 46,000 human milk fat globule protein that is highly expressed
 RT in human breast tumors contains factor VIII-like domains.";
 RL Cancer Res. 51:4994-4998(1991).
 RN [3]
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RC TISSUE-Milk;
 RX MEDLINE=98194924; PubMed=9535276;
 RA Giuffrida M.G., Cavalletto M., Giunta C., Conti A.,
 RA Godovac-Zimmermann J.;
 RT "Isolation and characterization of full and truncated forms of human
 RT breast carcinoma protein BA46 from human milk fat globule membranes.";
 RL J. Protein Chem. 17:143-148(1998).
 RN [4]
 RP SEQUENCE OF 268-317, AND IDENTIFICATION OF MEDIN.
 RX MEDLINE=99342076; PubMed=10411933;
 RA Haegqvist B., Naeslund J., Sletten K., Westermarck G.T., Mucchiano G.,
 RA Tjernberg L.O., Nordstedt C., Engstrom U., Westermarck P.;
 RT "Medin: an integral fragment of aortic smooth muscle cell-produced
 RT lactadherin forms the most common human amyloid.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:8669-8674(1999).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=97405885; PubMed=9260929;
 RA Taylor M.R., Couto J.R., Scallan C.D., Ceriani R.L., Peterson J.A.;
 RT "Lactadherin (formerly BA46), a membrane-associated glycoprotein
 RT expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp
 RT (RGD)-dependent cell adhesion.";
 RL DNA Cell Biol. 16:861-869(1997).
 RN [6]
 RP FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. BINDS
 CC SPECIFICALLY TO ROTAVIRUS AND INHIBITS ITS REPLICATION.
 CC [7]
 CC FUNCTION: MEDIN IS THE MAIN CONSTITUENT OF AORTIC MEDIAL AMYLOID.
 CC [8]
 CC SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
 CC [9]
 CC TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND AORTIC
 CC MEDIA. OVEREXPRESSED IN SEVERAL CARCINOMAS.
 CC [10]
 CC PTM: MEDIN HAS A RAGGED N-TERMINUS WITH MINOR SPECIES STARTING AT
 CC AMINO ACID 264 AND 273.
 CC [11]
 CC SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC [12]
 CC SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC [13]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC [14]
 CC EMBL, U58516; AAC50549.1; -;
 DR EMBL, S56151; AAB19771.1; -.

DR MM: 602281; -;
 DR InterPro: IPR000421; -;
 DR InterPro: IPR000561; -;
 DR Pfam: PF00008; EGF_1;
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 DR Signal: Glycoprotein; Milk; Repeat; EGF-like domain; Amyloid.
 FT SIGNAL 1 23
 FT CHAIN 24 387 LACTADHERIN.
 FT CHAIN 202 387 LACTADHERIN; SHORT FORM.
 FT CHAIN 268 317 MEDIN.
 FT DOMAIN 24 67 EGF-LIKE.
 FT DOMAIN 70 225 F5/8 TYPE C 1.
 FT DOMAIN 230 387 F5/8 TYPE C 2.
 FT SITE 46 48 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 27 38 BY SIMILARITY.
 FT DISULFID 27 38 BY SIMILARITY.
 FT DISULFID 32 55 BY SIMILARITY.
 FT DISULFID 57 66 BY SIMILARITY.
 FT DISULFID 70 225 BY SIMILARITY.
 FT DISULFID 212 216 BY SIMILARITY.
 FT DISULFID 230 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 387 AA; 43123 MW; 2BE6571DECB3782D CRC64;
 Query Match 38.1%; Score 1020.5; DB 1; Length 387;
 Best Local Similarity 48.8%; Pred. No. 2,7e-70;
 Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;
 QY 78 CIPNCHNGGTC-EISEAYRGDTFGVYCKCPRGNGIHCOHNI NECEVEPCKNGGICTD 136
 DB 27 CSKNCHNGGICEETSDQEVRGDPSTYCTCLKGAGH----- 65
 QY 137 LVANYSCECPGFEMGRNCGKSGPLGIEGIIISNOQTASSTHRALFGLOKWPYYARL 196
 DB 66 -----CETKVEPLGMEGNINIAQSIAASVAVTFGLQHWPELARL 108
 QY 197 NKKGLINMTAENDRMKRWIOINLQKKRYTGVTQGAKRGSPEYIKFKIANSNGK 256
 DB 109 NRAGVNAWTPSSND-NEWIQVNLRRMYGVVTOGASRLASHLEYLAKRVAVSLNGH 167
 QY 257 TW-AMKVKVGTEDWYFREGINDNTPYANSFPPIKAQVVRVLYPOVCRRHCTLRMELLGC 315
 DB 168 EFDFEIDVKKKHEFY--GNMKNAVHYNLFETPEAOYVRVLYPTSCHTACTLRRELLGC 225
 QY 316 ELSGSEPLGMSGHIQDYQITASSIFRTLNDMFTWEPKARLDKOGKVNAMTSGHNDQ 375
 DB 226 ELNGCANPLGLKNNISLPDQITASSYKTLWGLHPSMNSYARLDKOGFNAMVAGSGCN 285
 QY 376 SQMLQVLDLVPTKVTGIIITQGAKRGHQVQVSYKILANSDEHMYTQVODEQRDKYQ 435
 DB 286 DQMLQVLDLSSKKEVIGILITQGARNEFSQVQVASYKVAASNDASAMTETQDPRTGSKLTP 345
 QY 436 GNFQNDTHRRKNTDPTVYARHRIILPWSYGRITLASELLGC 477
 DB 346 GMDNDSHKKNLFETYLIRYVILIPVAMHNRIALRELLGC 387
 RESULT 6
 FA5_HUMAN STANDARD; PRT; 2224 AA.
 AC P12259; Q14285;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).
 GN F5.


```

FT SITE 1640 1641 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT MOD_RES 367 367 SUFFATATION (BY SIMILARITY).
FT MOD_RES 737 737 SUFFATATION (BY SIMILARITY).
FT MOD_RES 738 738 SUFFATATION (BY SIMILARITY).
FT MOD_RES 742 742 SUFFATATION (BY SIMILARITY).
FT MOD_RES 1669 1669 SUFFATATION (REQUIRED FOR VWF BINDING)
FT MOD_RES 1669 1669 (BY SIMILARITY).
FT MOD_RES 1687 1687 SUFFATATION (BY SIMILARITY).
FT MOD_RES 173 173 PROBABLE.
FT MOD_RES 547 573 PROBABLE.
FT MOD_RES 1819 1845 PROBABLE.
FT MOD_RES 2008 2156 BY SIMILARITY.
FT MOD_RES 2161 2313 BY SIMILARITY.
FT MOD_RES 61 61 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 233 233 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 259 259 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 423 423 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 601 601 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 880 880 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 958 958 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1015 1015 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1022 1022 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1026 1026 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1044 1044 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1076 1076 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1087 1087 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1136 1136 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1161 1161 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1192 1192 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1255 1255 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1268 1268 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1273 1273 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1274 1274 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1302 1302 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1316 1316 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1340 1340 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1378 1378 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 1797 1797 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 2105 2105 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT MOD_RES 2319 2319 AA: 266148 MW: PD054DE051DB2A01 CRC64:
SQ SEQUENCE

```

Query Match 25.0%; Score 670; DB 1; Length 2319;
 Best Local Similarity 46.3%; Pred. No. 7.1e-43;
 Matches 151; Conservative 40; Mismatches 119; Indels 16; Gaps 5;

```

QY 157 KCSGLGIEGSLNOCITASTSTRALFGLOKMPYATLAKGLINAWTAENDRKRW 216
DB 2007 QCOIFLGMASSIRDFQITASGHY-----GQWAPMLARLHSGSINAMSTKEP--FSW 2057
QY 217 IQINLQRMKRVYVITOGAKRIGSPKIKFYKTAISNDKRTAMVKVGTEDVFRGNI 276
DB 2058 IKVDLLAMIMHGIKTOGARQKFEISLISQFIMYSLDCKKMLSYQGSTGLWVFEQNV 2117
QY 277 DNNPTFANSFPPIKAOYVRLYPVCHRRHCTLRMELLSGELSGEPTLGMKSGHIOXOI 336
DB 2118 DSSGKHNHSPPIIARIKRLHPTHSIRSTLMELMGLDNLSCIPIDMESKYSIDQI 2177
QY 337 TASSIFRTLNDMF--TWPRKARLDKQKYNANTSGHNDOSWLOVDLVPKVTGIITQ 395
DB 2178 TASSYF---TNMFATWSPQARHLQGRTNARPOVNPCKWMLQVDLQKTKVYGIITQ 2233
QY 396 GAKDGHVGFVGSYKLAISNDGHEHMYODEKQKDKVQGNFNDTRKKNVIOPIYAR 455
DB 2234 GYKSLFTMFYKFEFLISSODGHHWT--QLLYNGKVKYVQGNDSSTPMANSLDPPLTR 2231
QY 456 HIRILPMSWYGRITLASELLGCTEE 481
DB 2292 YLRHPQIWEHQIALRLLEIGCEAQ 2317

```

RESULT 8
 FAS_BOVIN

```

ID FAS_BOVIN STANDARD: PRT: 2211 AA.
AC 028107: 028108;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).
GN F5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=liver;
RX MEDLINE=92147638; PubMed=1737753;
RA Gultio E.R., Esmen C.T., Mann R.G., Macgillivray R.T.;
RT "The complete cDNA sequence of bovine coagulation factor V.";
RL J. Biol. Chem. 267:2971-2978(1992).
CC -1 FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES
CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.
CC -1 SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT
CC CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS
CC IS CALCIUM-DEPENDENT.
CC -1 DOMAIN: DOMAIN B CONTAINS 29.5 X 9 AA TANDEM REPEATS, AND 2 X 17
CC AA REPEATS.
CC -1 PTM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE
CC COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-
CC TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).
CC -1 SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS. EACH IS COMPOSED OF
CC 2 PLASTOCYANIN-LIKE REPEATS.
CC -1 SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1 SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DR EMBL: M81440; AAA30512.1; -
DR EMBL: M81441; AAA30513.1; -
DR HSSP: P00450; 1KCW.
DR InterPro: IPR000421; -
DR InterPro: IPR001117; -
DR Pfam: PF00394; Cu-oxidase; 3.
DR Pfam: PF00754; F5_F8_Type-C; 2.
DR PROSITE: PS00079; MULTICOOPER_OXIDASE1; 2.
DR PROSITE: PS01285; FAS8C_1; 2.
DR PROSITE: PS01286; FAS8C_2; 2.
KW Blood coagulation; Plasma; Glycoprotein; Calcium; Signal; Zymogen;
KW Repeat.
FT SIGNAL 1 28
FT CHAIN 29 2211
FT CHAIN 29 741
FT PEPTIDE 742 1564
FT FT
FT CHAIN 1565 2211
FT DOMAIN 30 327
FT DOMAIN 30 193
FT DOMAIN 203 327
FT DOMAIN 348 686
FT DOMAIN 348 525
FT DOMAIN 348 686
FT DOMAIN 535 1564
FT DOMAIN 535 686
FT SIMILAR 696 915
FT DOMAIN 1124 1151
FT REPEAT 1124 1137
FT REPEAT 1138 1151
FT DOMAIN 1188 1453

```

POTENTIAL.
 COAGULATION FACTOR V
 HEAVY CHAIN (BY SIMILARITY).
 ACTIVATION PEPTIDE (CONNECTING REGION)
 (BY SIMILARITY).
 LIGHT CHAIN (BY SIMILARITY).
 F5/8 TYPE A 1.
 PLASTOCYANIN-LIKE 1.
 PLASTOCYANIN-LIKE 2.
 F5/8 TYPE A 2.
 PLASTOCYANIN-LIKE 3.
 PLASTOCYANIN-LIKE 4.
 B.
 TO 17 AA REPEATS IN HUMAN FAS.
 2 X 14 AA TANDEM REPEATS.
 1.
 2.
 30 X 9 AA TANDEM REPEATS OF [AS]-L-S-P-
 D-[LP]-[GS]-Q-[TE] (APPROXIMATE).

Query Match	Similarity	24.0%	Score 657.5	DB 1	Length 2211
Best Local	Similarity	42.0%	Pred. No. 6e-42		
Matches 136	Conservative	56	Mismatches 121	Indels 11	Gaps 4
Oy	157	KCSGPGICBEGIIISNOQITASTSTHRALEFLOKMTPTPYARLNKGLINAWTA--AENDRW	213		
Db	1893	ECMPMGISLGLIADSOIADSE---FNGY--WEPKLARLNNGSVYAWTAETLSTFNP	1946		
Oy	214	KRMTOJNLORKMVTSTVITOGAKRIRGSEPEYIKFKIAYSDNGKTMAYKYKGTNEDMVR	273		
Db	1947	EPPIQVDMQKEVLLIIOIQGAKHLKRYTTEFCVATSIDRKMMRIEKENSTRNMYG	2006		
Oy	274	GNIDNNTPYANSTFTPIKQVRLPYOYCRHCHTRMELLACELSGCSSEPLGKSGHID	333		
Db	2007	GNDASTIKENQIDPEPVARYIRISPTGSYNKPALRIETLQCEVNGCSTPLGMESEKIE	2066		
Oy	334	YQTSASITPELTLMQDFTEPPKRALIDQKQKYNMTSGNDOSQMLQVDLVTPTKVTGI	393		
Db	2067	KQITASSFFKSMWGN--WEPLARLAINQGRNMAQAANNQWQIDDLKIKITAIY	2124		
Oy	394	TQCAKDFGHQYFVGSKTLAYSNDGEHVTYVODEKOKKRYQCGFNDTRKNVIDPPI	453		
Db	2125	TQCGKSLSEMYKYSTYIHSDQFTDMKPYREKSSMDKIFEGNNVNGHVKNFENPPI	2184		
Oy	454	ARHRIPLMSWYGRITLASELLGC	477		
Db	2185	SFRIRIPIKTNOSIALRIELFGC	2208		
RESULT	9				
FA8_HUMAN	STANDARD	PRT	2351	AA	
AC	P00451				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	COAGULATION FACTOR VIII PRECURSOR (PROCOCAGULANT COMPONENT)				
GN	(ANTHEMOPLHILIC FACTOR) (AHF).				
GN	FR OR FRC.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
LN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=86081164; PubMed=3935400;				
RA	Treut M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,				
RA	Hartig C., Kuo C.H., Maslarz F.R., Merriweather J.P., Najarian A.,				
RA	Rachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,				
RA	Urdea M.S., Valenzuela P., Dahl H.-H.M., Favaloro J., Hansen J.,				
RT	Nordfang O., Ezban M.;				
RT	"Characterization of the polypeptide composition of human factor				
RT	VIII.C and the nucleotide sequence and expression of the human kidney				
RT	cDNA.";				
RL	DNA 4:333-349(1985).				
LN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85061548; PubMed=6438526;				
RA	Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschler J.,				
RA	Keit B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,				
RA	Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.W.;				
RT	"Expression of active human factor VIII from recombinant DNA clones.";				
RT	Nature 312:330-337(1984).				
LN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85061550; PubMed=6438528;				
RA	Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,				
RA	Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,				
RA	Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Pass D.N.,				
RA	Heppick R.M.;				
RT	"Molecular Cloning of a cDNA encoding human antithrombophilic factor.";				
RT	Nature 312:342-347(1984).				

RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93265012; PubMed=1303178;
 RA Gitschler J., Wood W.I.;
 RT "Sequence of the exon-containing regions of the human factor VIII
 gene";
 RL Hum. Mol. Genet. 1:199-200(1992).
 RN [5]
 RP SEQUENCE OF 2064-2070 FROM N.A.
 RA de Water N.S., Williams R., Browett P.J.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SULFATATION OF 1699.
 RX MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Nlehrs C., Huttner W.B., Verbeet M.P.,
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 essential for the interaction of factor VIII with von Willebrand
 factor";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RP SULFATATION.
 RX MEDLINE=92207952; PubMed=1554716;
 RA Pitman D.D., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate
 residues within recombinant factor VIII";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX MEDLINE=95200924; PubMed=7893714;
 RA Gilbert G.E., Baleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 amphipathic structure as determined by NMR spectroscopy";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=91221499; PubMed=1902642;
 RA Gitschler J.;
 RT "The molecular basis of hemophilia A";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT GLN-2326.
 RX MEDLINE=86235434; PubMed=3012775;
 RA Gitschler J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 mild hemophilic";
 RL Science 232:1415-1416(1986).
 RN [13]
 RP VARIANT PRO-2135.
 RX MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschler J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 analysis of amplified hemophilia DNA sequences";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT GLN-2228.
 RX MEDLINE=88191889; PubMed=2833855;
 RA Yousoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 relative mutation rate at CG dinucleotides";

RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT GLY-291.
 RX MEDLINE=88220354; PubMed=2835904;
 RA Yousoufian H., Wong C., Aronis S., Plakoukakis H., Kazazian H.H. Jr.,
 RT Antonarakis S.E.;
 RL "Moderately severe hemophilia A resulting from Glu->Gly substitution
 in exon 7 of the factor VIII gene";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT CYS-1708.
 RX MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Purification and characterization of factor VIII 1,689-Cys: a
 nonfunctional cofactor occurring in a patient with severe hemophilia
 A";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RP VARIANT CYS-391.
 RX MEDLINE=90001343; PubMed=2506948;
 RA Shima M., Ware J., Yoshioke A., Fukui H., Fulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 thrombin cleavage site in a dysfunctional factor VIII molecule";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RP VARIANT LEU-189.
 RX MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 resulting in moderately severe hemophilia A";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RP VARIANT LEU-2326.
 RX MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 of the factor VIII gene";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RP VARIANT HIS-391.
 RX MEDLINE=89264602; PubMed=2498882;
 RA Araki M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RT Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 mutation altering a thrombin cleavage site
 (arginine-372->histidine)";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT CYS-1708.
 RX MEDLINE=90105723; PubMed=2104766;
 RA Araki M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RT Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 Cys) in the factor VIII gene of two unrelated patients with
 cross-reacting material positive hemophilia A";
 RL Blood 75:384-389(1990).
 RN [22]
 RP VARIANTS GLN-2228 AND LEU-2326.
 RX MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RT Mancuso G., Morfini M., de Biasi R., Bando F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 VIII gene of hemophilia A patients of Italian descent";
 RL Blood 75:662-670(1990).
 RN [23]
 RP VARIANT CYS-391.
 RX MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Afani A., Tuddenham E.G.;
 RT "CMT haemophilia A due to a missense mutation (372->Cys) at the
 internal heavy chain thrombin cleavage site";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RP VARIANTS PHE-1699 AND CYS-1708.

RX MEDLINE=90152691; PubMed=2105906;
RA Higuchi M., Wong C., Kochan L., Olek K., Aronis S., Kasper C.K.,
RA Kazanian H.H., Antonarakis S.E.,
RT "characterization of mutations in the factor VIII gene by direct
RT sequencing of amplified genomic DNA.",
RL Genomics 6:65-71(1990).
RN [25]
RN VARIANTS CYS-1728 AND ASP-1941.
RP MEDLINE=90169988; PubMed=2106480;
RX Travistman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
RA Kazanian H.H.,
RT "use of denaturing gradient gel electrophoresis to detect point
RT mutations in the factor VIII gene.";

Query Match	24.3%	Score 649;	DB 1;	Length 2351;
Best Local Similarity	44.28;	Pred. No. 2,98-41;		
Matches 144;	Conservative	46;	Mismatches 120;	Indels 16;
			Gaps	5
QY 157	KCSGPGICIEGIIISNOOITASSTHRALBELQKWYYVARLNKKLIINAWTAENDRMKRW	216		
DB 2039	KCQPIPLGMSAGHRIIDPOIRASGQY-----GGMAPKILRLHYSGINAWMSKPEP---FSW	2089		
QY 217	IQIILQKKHNVTVITQIGAKRIGSPYEIKFYIAYSDNGKTMAYKVKGTINEDVPRGNI	276		
DB 2090	IKVDLLAPMLIHRIKIQGAKRQKRFSSLYISQFIIMSLDGKKMQYTRGNSGTGLMVFPGNW	2149		
QY 277	DNMTPVANSFPPRIKAQVRYRLPYQVCRHRCTLRMELLCGELSGCSEPIGKSGHIDQYI	336		
DB 2150	DSSGIRKINIFNPPIIARVYRLHPHTHYSIKSTLRMELMCGDINCSCSMPLGMSKAI	2509		
QY 337	TASSIETPLMMDF--TWEPKARLRQKQGVNVAWMTSGHNDQSMILOVLLVPYTKVGI	395		
DB 2210	TASSYF-----TNNFAWPSKAKLHLQGRSNMARPQVNNPKEMQLQVDPQKTMATG	2265		
QY 396	GAKPGFVQVGSYKAYASNDGSEHMTVYDDEQQRDKYAFQCNPFNDTHRRKNVDDPIYAR	455		
DB 2266	GVSKELLTSMYVKRELLISSQDGHQWTLF--PQNGVVKYKFGQNDQSPFPYVNASIDPPLTR	2323		
QY 456	HIRLIPMSWGRITLASELCTGTEEE	481		
DB 2324	YLRIHQSWVHOIALREVEVLGECAOD	2349		

```

RESULT 10
F8B_PIG ID F8B_PIG STANDARD PRI: 2133 AA.
AC P12263; Q95243;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT).
CN C9B.
OS Sus. scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Healey J.F., Lubin I.M., Lollar P.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RX SEQUENCE OF 705-1573 FROM N.A.
RX MEDLINE=86287369; PubMed=3016730;
RA Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
RA Kaufman R.J.;
RT "A large region (approximately equal to 95 kDa) of human factor VIII
RT is dispensable for in vitro procoagulant activity.";
RT Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
RN [3]
RP SEQUENCE OF 392-759 FROM N.A.
RX MEDLINE=94179260; PubMed=7510693;
RA Healey J.F., Scandella D., Runge M.S., Lollar P.;
RT "Elimination of a major inhibitor epitope in factor VIII.";

```

RL	J. Biol Chem.	269:8639-8641(1994).		
CC	-I- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE ACTIVATED FORM, FACTOR XA.			
CC	-I- SUBCELLULAR LOCATION: EXTRACELLULAR.			
CC	-I- SIMILARITY: CONTAINS 3 F5/F8 TYPE A DOMAINS; EACH IS COMPOSED OF 2 PLASTOCYANIN-LIKE REPEATS.			
CC	-I- SIMILARITY: CONTAINS 2 F5/F8 TYPE C DOMAINS.			
CC	-I- SIMILARITY: STRONG, TO COAGULATION FACTOR V.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/or send an email to licensel@isb.slb.ch).			
CC	----			
DR	EMBL; U49517; AAB06705.1; .			
DR	PfR; A25945; A25945.			
DR	HSSP; P00451; ICFG.			
DR	InterPro; IPR000421; .			
DR	InterPro; IPR001117; .			
DR	Pfam; PF00734; Cu-oxidase; 3.			
DR	Pfam; PF00754; F5_F8_Type_C_2.			
DR	PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.			
DR	PROSITE; PS01285; FASRC_1; 2.			
DR	PROSITE; PS01286; FASRC_2; 2.			
KW	Blood coagulation; Repeat; Plasma; Acute phase; Calcium; Signal; glycoprotein; Sulfatation.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	2133	COAGULATION FACTOR VIII.
FT	DOMAIN	20	357	F5/F8 TYPE A 1.
FT	DOMAIN	20	199	PLASTOCYANIN-LIKE 1.
FT	DOMAIN	207	357	PLASTOCYANIN-LIKE 2.
FT	DOMAIN	399	730	F5/F8 TYPE A 2.
FT	DOMAIN	399	573	PLASTOCYANIN-LIKE 3.
FT	DOMAIN	583	730	PLASTOCYANIN-LIKE 4.
FT	DOMAIN	760	1599	B.
FT	DOMAIN	1495	1822	F5/F8 TYPE A 3.
FT	DOMAIN	1495	1659	PLASTOCYANIN-LIKE 5.
FT	DOMAIN	1669	1822	PLASTOCYANIN-LIKE 6.
FT	DOMAIN	1822	1970	F5/F8 TYPE C 1.
FT	DOMAIN	1975	2127	F5/F8 TYPE C 2.
FT	SITE	391	392	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	759	760	CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT	SITE	1449	1450	CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT	SITE	1490	1491	CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT	MOD_RES	737	737	SULFATATION (BY SIMILARITY).
FT	MOD_RES	738	738	SULFATATION (BY SIMILARITY).
FT	MOD_RES	742	742	SULFATATION (BY SIMILARITY).
FT	DISULFD	173	199	SULFATATION (BY SIMILARITY).
FT	DISULFD	547	573	PROBABLE.
FT	DISULFD	1633	1659	PROBABLE.
FT	DISULFD	1822	1970	PROBABLE.
FT	DISULFD	1975	2127	PROBABLE.
FT	CARBOHYD	233	233	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	601	601	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	929	929	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	985	985	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	929	929	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1025	1025	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1111	1111	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1181	1181	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1208	1208	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1245	1245	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1265	1265	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1335	1335	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1408	1408	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1611	1611	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1919	1919	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CONFLICT	713	713	N -> M (IN REF. 2).
FT	CONFLICT	734	734	I -> T (IN REF. 2).

FT CONFLICT 792 792 G -> Q (IN REF. 2).
 FT CONFLICT 1133 1133 E -> F (IN REF. 2).
 FT CONFLICT 1191 1191 I -> L (IN REF. 2).
 FT CONFLICT 1209 1209 R -> F (IN REF. 2).
 FT CONFLICT 1437 1437 C -> G (IN REF. 2).
 FT CONFLICT 1456 1456 F -> R (IN REF. 2).
 FT CONFLICT 1539 1539 F -> R (IN REF. 2).
 FT CONFLICT 1546 1546 O -> N (IN REF. 2).
 SO SEQUENCE 2133 AA, 239304 MW, 152BBA8997570DA CRC64;

Query Match 23.9%; Score 639; DB 1; Length 2133;
 Best Local Similarity 40.0%; Pced. No. 1.5e-40;
 Matches 144; Conservative 52; Mismatches 136; Indels 28; Gaps 6;

QY 122 ECEVPCKNKGICITDLVANSCECPGFEMGRNCQYKSGFLIEGGIISNOQITASSYHR 181
 DB 1800 ECLIGEHLOAGMSTTFLV-YSKR-----CQAPLGMASGRIRDQITASQY- 1844
 QY 182 ALFGLQWYPPYARLNKKGLINMTAENDRMKRWIQLNLRKRVGVTGQAKRIGSP 241
 DB 1845 -----GQWAPKLARLHYSGSINAMSTRDP---HSWIKYDLAPMTHIGIMTOGAROFSS 1896
 QY 242 EYKFKYKAVSNDGKTAMKYKGVGNDVYFNGNDNNPNVANSPTPIKAQYRLKPOV 301
 DB 1897 LYISGFITMYSLDGRNMQYRGNSGTGLAVFGVNDVASGKINIFNPPIVARIIRLHPTH 1956
 QY 302 CRRHCTLMELGCELSGCEPLGKSGHIDQYQTASSIFPTLMKMFWEPRKARLDK 361
 DB 1957 YSIRSTLEMELMGCLNSCMPGDMONKALSDQITASS--HLSNIRATWSPQARHL 2013
 QY 362 QGKVANATSGHNDQSQMLQYDLVPTKVGITIGOGAKDGHQVFGSKYKLVANSNGEHW 421
 DB 2014 QGRTANAMPKRVSSAEMWLYQDLQKTVKVGITIGOGKVLSSMYKKEFLVSSQDGRMT 2073
 QY 422 VYQDQKQKXDVYQGNFNDTHKKNVDPPIVARIIRLHPTMWSGRITLASELIGCTEE 481
 DB 2074 LFLQDGH--TKVFGQNDQSTPYVVALDPLFLTRIRHPTMAOHIALRLVLEACEAD 2131

RESULT 11
 NRPI_HUMAN STANDARD; PRT; 923 AA.
 AC 014786; 060461;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROPLIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM).
 RA MEDLINE-97435084; PubMed-9288753;
 RA He Z., Tessier-Lavigne M.;
 RT "Neuropilin is a receptor for the axonal chemorepellent semaphorin III.";
 RL Cell 90:739-751(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM), AND SEQUENCE OF 22-39.
 RA TISSUE-Breast;
 RA MEDLINE-96186099; PubMed-95292350;
 RA Soker S., Takashima S., Miao H.-Q., Neufeld G., Klagsbrun M.;
 RT "Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor.";
 RL Cell 92:735-745(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (SOLUBLE/SNRPI ISOFORM), AND SEQUENCE OF 22-31.
 RA TISSUE-Prostatic adenocarcinoma;
 RC MEDLINE-20183929; PubMed-10688860;

RA Gagnon M.L., Bielenberg D.R., Gechtman Z., Miao H.-Q., Takashima S.,
 RA Soker S., Klagsbrun M.;
 RT "Identification of a natural soluble neuropilin-1 that binds vascular
 RT endothelial growth factor: in vivo expression and antitumor
 RT activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2573-2578(2000).
 RN [4]
 RP CHARACTERIZATION.
 RA MEDLINE-20309748; PubMed-10748121;
 RT Gluzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
 RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid
 RT form of vascular endothelial growth factor (VEGF) and of placenta
 RT growth factor-2, but only neuropilin-2 functions as a receptor for
 RT the 145-amino acid form of VEGF.";
 RL J. Biol. Chem. 275:18040-18045(2000).
 CC -1- FUNCTION: THE MEMBRANE-BOUND ISOFORM IS A RECEPTOR INVOLVED IN THE
 CC DEVELOPMENT OF CERTAIN NEURONAL CIRCUITS AND IN ANGIOGENESIS. IN THE
 CC FORMATION OF CERTAIN NEURONAL CIRCUITS AND IN ANGIOGENESIS
 CC OUTSIDE THE NERVOUS SYSTEM. IT MEDIATES THE CHEMOREPULSANT
 CC ACTIVITY OF SEMAPHORINS. IT BINDS TO SEMAPHORIN 3A, THE PLGF-2
 CC ISOFORM OF PGF, THE VEGF-165 ISOFORM OF VEGF AND VEGF-B.
 CC COEXPRESSION WITH KDR RESULTS IN INCREASED VEGF-165 BINDING TO KDR
 CC AS WELL AS INCREASED CHEMOTAXIS. IT MAY REGULATE VEGF-INDUCED
 CC ANGIOGENESIS.
 CC -1- FUNCTION: THE SOLUBLE/SNRPI ISOFORM BINDS VEGF-165 AND APPEARS
 CC TO INHIBIT ITS BINDING TO CELLS. IT MAY ALSO INDUCE APOPTOSIS BY
 CC SEQUESTERING VEGF-165. MAY BIND AS WELL VARIOUS MEMBERS OF THE
 CC SEMAPHORIN FAMILY. ITS EXPRESSION HAS AN AVERSE EFFECT ON BLOOD
 CC VESSEL NUMBER AND INTEGRITY.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE SOLUBLE/SNRPI
 CC ISOFORM IS SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; MEMBRANE-BOUND (SHOWN HERE)
 CC AND SOLUBLE/SNRPI; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: THE EXPRESSION OF MEMBRANE-BOUND AND
 CC SOLUBLE/SNRPI ISOFORMS DOES NOT SEEM TO OVERLAP. THE MEMBRANE-
 CC BOUND ISOFORM IS EXPRESSED BY THE BLOOD VESSELS OF DIFFERENT
 CC TISSUES IN THE DEVELOPING EMBRYO IT IS FOUND PREDOMINANTLY IN THE
 CC NERVOUS SYSTEM. IN ADULT TISSUES, IT IS HIGHLY EXPRESSED IN HEART
 CC AND PLACENTA; MODERATELY IN LUNG, LIVER, SKELETAL MUSCLE, KIDNEY
 CC FOUND IN LIVER HEPATOCYTES, KIDNEY DISTAL AND PROXIMAL TUBULES.
 CC -1- SIMILARITY: BELONGS TO THE NEUROPLIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 DR EMBL: AF018946; AAC51759.1; -;
 DR EMBL: AF016050; AAC12921.1; -;
 DR EMBL: AF145712; AAF44344.1; -;
 DR MIM: 602069; -;
 DR InterPro: IPR000421; -;
 DR InterPro: IPR000859; -;
 DR InterPro: IPR000988; -;
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00754; F5_P8_Type_C; 2.
 DR Pfam: PF00629; MAM; 1.
 DR PRINTS: PR00020; MAMDOMAIN.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 DR PROSITE: PS00740; MAM_1; 1.
 DR PROSITE: PS0060; MAM_2; 1.
 DR Transmembrane: Glycoprotein; Neurone; Signal; Repeat; Receptor;
 KW Alternative splicing;
 FT SIGNAL 1 21

```

FT CHAIN 22 923 NEUROFILIN-1.
FT DOMAIN 22 856 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 857 879 POTENTIAL.
FT DOMAIN 880 923 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 141 CUB 1.
FT DOMAIN 147 265 CUB 2.
FT DOMAIN 275 424 F5/8 TYPE C 1.
FT DOMAIN 431 583 F5/8 TYPE C 2.
FT DOMAIN 645 811 MAM.
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 642 644 EEP -> GIK (IN SOLUBLE/SNRP1 ISOFORM).
FT VARSPLIC 645 923 MISSING (IN SOLUBLE/SNRP1 ISOFORM).
FT CONFLICT 26 26 K -> E (IN REF. 1).
FT CONFLICT 749 749 D -> H (IN REF. 2).
FT CONFLICT 855 855 E -> D (IN REF. 2).
SQ SEQUENCE 923 AA; 103120 MW; ADEADC4A849E5D57 CRC64;

```

```

Query Match 17.4%; Score 465.5; DB 1; Length 923;
Best Local Similarity 32.1%; Pred. No. 7.9e-28;
Matches 159; Conservative 59; Mismatches 210; Indels 67; Gaps 23;

```

```

QY 12 GLSLGVPOFGSGDIDCPN-PCENGGICLPGLAVG-SFSCPCPGFDPDCCSSV-EVAVSD 68
DB 13 GFSITVELFKKGPSPCSQNTTPPSVIVKSPGFPKRYPNSECTIYVAPKMSILLIEFESF 192
QY 69 EEPYTSAGCPTPNCPCHNGTCEISEAYRGDTF-----IGYVC--KCP--RGFNGICHO 117
DB 193 DLEPDSNP-----GGMFCRYDRLEIMDGFDPVGPPIGRVCGQKTPGRIRSSSGILSM 245
QY 118 HNINCEVEPKNGGICGICDILVANYSC--ECRGEFMGRNQYKCSGPLGEGGILISNOQT 175
DB 246 VFYIDSAI--AKEG-----FSANISVLQSSVSF-----DFKCMELMGSEGEIHSDIT 292
QY 176 ASSTRALFGLQKWPPYARLNKGLINAMTAENDRMKRWIDINLQRMKRYGVITOGA 235
DB 293 ASSQYST-----NMSAERSRLNYPE--NGWTPGE-DSYREMIQVDGLIRFYTAGTOGA 344
QY 236 --KRIGSPEYIKFKYIAVNDGKTWAMYKYKGTNEDMVRFGINDNTPTVANSFETPIKAQ 293
DB 345 ISKETKKKYYIKYKIDYSSNGEDM--ITTEKGNKPYLPQGNTPIDVYVAVPEKPLITR 402
QY 294 YVRLYPOVCRRHCTLRMELLGCELSSG--CSEPLGMSGHIQDYQITASSIFETLNMDET 351
DB 403 EVRIKPAWTEGISMREFYVCGKITDYPCSGMLGWSGLISDSQITSS-----NOGDEN 456
QY 352 WEPRKARL--DKQCKVAMT-----SGHNDQSONLOVDLVPYTKYGIITOGAKDFHVGQV 406
DB 457 WMPENIRLVYTSRG---WALPPAPHSYINEMIQIDIGEEKLYRGILITOGGHHREKKYFM 512
QY 407 GSYKLAVNDEHMTVYODEKORAKVFOGNFNDTRHRKNVVIDPPIYAHRIPL-WSNY 465
DB 513 RKFKIGYGNNSDMKIMINDSKRAKSEEGNNNVDPELRTF-PALSTPFIIRIYERATH 571
QY 466 GRITIASLGLCTEE 480
DB 572 GGLGLRMELLGCEVE 586

```

```

RESULT 12
NRPL_CHICK STANDARD; PRT; 914 AA.
AC P79793;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN).
GN NRPL OR NRPL.
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=Embryonic brain;
RX MEDLINE=95324761; PubMed=7601310;
RA Takagi S., Kasuya Y., Shimizu M., Matsuura T., Tsudoi M., Kawakami A.,
RJ Fujisawa H.;
RT *Expression of a cell adhesion molecule, neuropilin, in the
RL developing chick nervous system.*;
RL Dev. Biol. 170:207-222(1995).
CC -1- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CC CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CC CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY
CC SIMILARITY). SEEMS TO HAVE CALCIUM-INDEPENDENT CELL ADHESION
CC PROPERTIES.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: DEVELOPING NERVOUS SYSTEM: OPTIC TECTUM
CC (LAYERS D AND E OF SGFS), AMACRINE CELLS OF RETINA, NEURITES OF
CC DORSAL ROOT GANGLIA. ALSO EXPRESSED IN NONNEURONAL CELLS, E.G.
CC BLOOD VESSELS BELONGS TO THE ENTIRE EMBRYO.
CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D45416; BAA08256.1; -
DR InterPro; IPR000421; -
DR InterPro; IPR000859; -
DR InterPro; IPR000998; -
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00754; F5_F8_Type_C; 2.
DR Pfam; PF00629; MAM; 1.
DR PRINTS; PRO0020; MAMDOMAIN.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS50060; MAM_2; 1.
KW Transmembrane; Glycoprotein; Neutone; Signal; Repeat; Receptor;
KW Cell adhesion.
FT CHAIN 1 18
FT SIGNAL 1 914
FT DOMAIN 20 847
FT TRANSEM 848 870
FT DOMAIN 871 914
FT DOMAIN 925 139
FT DOMAIN 145 263
FT DOMAIN 273 422
FT DOMAIN 429 581
FT DOMAIN 636 801
FT DISULFID 25 52
FT DISULFID 80 102
FT DISULFID 145 171
FT DISULFID 204 226
FT DISULFID 273 422

```

FT DISULFID 429 581 BY SIMILARITY
SQ SEQUENCE 914 AA: 102480 MM; DDEE6D6F0CBB68C CRC64;

Query Match 17.3%; Score 461.5; DB 1; Length 914;
Best Local Similarity 31.6%; Pred. No. 1.6e-27;
Matches 156; Conservative 63; Mismatches 211; Indels 63; Gaps 23;

12 GLSLGVPOFGKGDICDPN-PCENGICLPGIAVSG-ECDFGDPDPCNGSVV-EVAD 68
131 GFSIRYEVFRKPECSNFTSSGMIKSPGFEKYPNLSLEYITIFAPKSEIIEESF 190
69 EEEPTAGPCTPNCHNGGTCEISEAYRGDTF-----IGYVC--KCP--RGFNGIHCQ 117
191 ELEPDSNTP-----GGAFCRDRLEIMDGPPEVPHIGRYCGQNNNGVRSSGTILSM 243
118 HINICEVEPCNGKIGICTDLVANYSCDFEPMGRNCOYKCSGPIGIGIISNOQITAS 177
244 VFYDSDAI--AKEG-----FSANYSVQSSV-----SEDFQCEPLGMSGEIHSQDITVS 292
178 STRHALFGLQKWPYARLNKGLINAWTAENDRMKRWIOINLQRMKRVTVITGGA-- 235
293 SOYSAI-----WSERSRLNYPE--NGWTPGE-DSVRMIDVLDGLLRFVSGIGTQGAIS 344
236 KRISPEYIKYKIAVSNDOCKTAMKVKYGTEDVAFRGNTDNNTPYANSTPPIKAOYV 295
345 KETKEKEYLKTIVDVSSNGEDM--TLLEKSGKPYVFOGNSNPDIYVRPEKPVLTIRFV 402
296 RLVPQVRRICITLMEELGCEISG--CSEPLGMSGHIQDIQTASSITFTLNKDMFTWE 353
403 RIKFVSNWENGVSIRFEYVCKITIDYPCSGMLGMSGLIPDSQITAST-----QVDR-NWI 456
354 PRKARL--DRQGVNANMT---SGHNDOSQILOVDLIVPTKVTYITGACPFCHVQFVS 408
457 PENARLITSSSG-----WALPPTHTPTNEMLOIDLEEKIYRGIIYVGKREKAYVMK 512
409 YKLAISMDGHWTVYODEKQKDYFOGFDNDTIRKKNVIDPPIYARHILRP--WSNYGR 467
513 FKIGYSNNGSDMKIMDSKIKIKIEGNTNYDTELRFE--PVSTRIRIVPERAVHAG 571
468 ITLASELIGCTEE 480
572 LGLRMEELGCELE 584

RESULT 13
NRPI_MOUSE STANDARD; PRT; 923 AA.
AC P97333;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN).
GN NRPI OR NRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Embryonic brain;
RX MEDLINE=96353149; PubMed=8748368;
RA Kawakami A., Kitsuikawa T., Takagi S., Fujisawa H.;
RT "Developmentally regulated expression of a cell surface protein,
neurofilin, in the mouse nervous system.";
RL J. Neurobiol. 29:1-17(1996).
CC -!- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
CERTAIN NEURONAL CIRCUITS AND IN ORGANOMERIS OUTSIDE THE NERVOUS
SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS. IT
BINDS TO SEMAPHORIN 3A, THE PLGF-2 ISOFORM OF PGF, THE VEGF-165
ISOFORM OF VEGF AND VEGF-B. COEXPRESSION WITH KDR RESULTS IN
INCREASED VEGF-165 BINDING TO KDR AS WELL AS INCREASED CHEMOTAXIS.

CC IT MAY REGULATE VEGF-INDUCED ANGIOGENESIS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: NERVOUS SYSTEM.
CC -!- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb.ch).
CC -----
DR EMBL: D50086; BAA08789.1; -
DR MCD: MG1:106206; NRP.
DR InterPro: IPR000421; -
DR InterPro: IPR000859; -
DR InterPro: IPR000998; -
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00754; F5_F8_Type_C; 2.
DR Pfam: PF00629; MAM; 1.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS01285; FAS8C_1; 2.
DR PROSITE: PS01286; FAS8C_2; 2.
DR PROSITE: PS00740; MAM_1; 1.
DR PROSITE: PS50060; MAM_2; 1.
DR Transmembrane; Glycoprotein;
KW SIGNAL 1 21
FT CHAIN 22 923
FT DOMAIN 22 856
FT TRANSMEM 857 879
FT DOMAIN 880 923
FT DONAIN 27 141
FT DONAIN 147 265
FT DONAIN 275 424
FT DONAIN 431 583
FT DONAIN 645 811
FT DONAIN 82 54
FT DISULFID 27 54
FT DISULFID 82 104
FT DISULFID 147 173
FT DISULFID 206 228
FT DISULFID 275 424
FT DISULFID 431 583
FT CARBOHYD 150 150
FT CARBOHYD 261 261
FT CARBOHYD 300 300
FT CARBOHYD 522 522
FT CARBOHYD 842 842
SO SEQUENCE 923 AA; 103020 MM; 0644B8A170796808 CRC64;

Query Match 17.1%; Score 458.5; DB 1; Length 923;
Best Local Similarity 31.9%; Pred. No. 2.7e-27;
Matches 158; Conservative 58; Mismatches 212; Indels 67; Gaps 23;

12 GLSLGVPOFGKGDICDPN-PCENGICLPGIAVSGSC-ECDFGDPDPCNGSVV-EVAD 68
131 GFSIRYEVFRKPECSNFTSSGMIKSPGFEKYPNLSLEYITIFAPKSEIIEESF 192
69 EEEPTAGPCTPNCHNGGTCEISEAYRGDTF-----IGYVC--KCP--RGFNGIHCQ 117
193 DLEQDSNTP-----GGMFCRDRLEIMDGPPEVPHIGRYCGGKTGRISSGVLGM 245
118 HINICEVEPCNGKIGICTDLVANYSC--ECDFGPMGRNCOYKCSGPIGIGIISNOQIT 175
246 VFYDSDAI--AKEG-----FSANYSVQSSV-----SEDFQCEPLGMSGEIHSQDIT 292
176 ASSTRHALFGLQKWPYARLNKGLINAWTAENDRMKRWIOINLQRMKRVTVITGGA 235
293 ASOYGT-----NMSVRSRLNYPE--NGWTPGE-DSYKEMIDVLDGLLRFVTVAVTGA 344

GenCore version 4.5
Copyright (c) 1993 - 2000. Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:20:42 ; Search time 105.14 Seconds
(without alignments)
536.209 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513
Perfect score: 2675
Sequence: 1 MKRSVAVMLVGLSLGVPOR.....MSWKGRTTLASLLGCTEEB 481

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTEMBL_15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2633.5	98.4	480	1	043854 homo sapien
2	2529.5	94.6	480	11	035474 mus musculu
3	1177	44.0	463	11	09R1X9 09R1X9 mus musculu
4	1154.5	43.2	426	11	09WTS3 09WTS3 mus musculu
5	972	36.3	363	6	077718 equus caball
6	682.5	25.5	2183	11	088783 mus musculu
7	670.5	25.1	2224	4	043737 homo sapien
8	648	24.2	2343	6	062730 canis fami
9	644	24.1	2343	6	018806 canis fami
10	477	17.8	858	5	076470 lytechinus
11	464.5	17.4	216	4	014286 homo sapien
12	455.5	17.0	779	4	09NTT3 09NTT3 homo sapien
13	447.5	16.7	921	11	09QX38 09QX38 rattus norv
14	428.5	16.0	192	11	09QWQ1 09QWQ1 mus musculu
15	330	12.3	2352	5	061240 halocynthia
16	327.5	12.2	3644	5	09VU94 drosophila
17	327.5	12.2	3843	5	09U5D0 drosophila
18	318	11.9	406	5	025059 heliocardia
19	316.5	11.8	2653	5	025253 lucilia cup

20	314	11.7	2447	13	013149	013149 fugu rubrip
21	308.5	11.5	529	5	025058	025058 heliocardia
22	306	11.4	585	5	09U0E2	09U0E2 tribolium c
23	302.5	11.3	2634	5	09W4F8	09W4F8 drosophila
24	302.5	11.3	2704	5	097458	097458 drosophila
25	301.5	11.3	752	13	042374	042374 brachydanio
26	297	11.1	1203	11	006008	006008 mus musculu
27	297	11.1	2470	11	035516	035516 lytechinus
28	297	11.1	2531	5	016004	016004 lytechinus
29	296	11.1	1471	11	09QW30	09QW30 rattus sp.
30	292.5	10.9	1404	5	09V855	09V855 drosophila
31	291.5	10.9	2319	11	09R172	09R172 rattus norv
32	288.5	10.8	2281	4	09UPL3	09UPL3 homo sapien
33	288.5	10.8	2321	4	09Y6L8	09Y6L8 homo sapien
34	288.5	10.8	2321	4	09UM47	09UM47 homo sapien
35	288	10.8	1530	11	09WUG5	09WUG5 rattus norv
36	287.5	10.7	589	11	088671	088671 rattus norv
37	287.5	10.7	615	13	057409	057409 brachydanio
38	287	10.7	717	13	P87357	P87357 brachydanio
39	287	10.7	723	4	09UJV2	09UJV2 homo sapien
40	287	10.7	723	4	09NU41	09NU41 homo sapien
41	286.5	10.7	156	5	026661	026661 strongyloce
42	286.5	10.7	802	13	057462	057462 brachydanio
43	285.5	10.7	721	13	091902	091902 xenopus lae
44	285.5	10.7	728	13	090656	090656 gallus gall
45	283.5	10.6	1193	13	090819	090819 gallus gall

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRF:	480 AA.
ID	043854	043854:		
AC	043854	PRELIMINARY:	PRF:	480 AA.
DT	01-JUN-1998	(FIREMBLrel. 06, Created)		
DT	01-JUN-1998	(FIREMBLrel. 06, Last sequence update)		
DT	01-OCT-2000	(FIREMBLrel. 15, Last annotation update)		
DE	INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.			
OS	DELL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.			
RC	TISSUE-EMBRYONIC LUNG;			
RX	MEDLINE=98083109; PubMed=9420328;			
RA	Hidai C., Zupancic T.J., Penta K., Mikhail A., Kawana M.,			
RA	Quertemous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,			
RA	Auerbach R., Hogan B.L.M., Snodgrass R., Quertemous T.;			
RT	"Cloning and characterization of developmental endothelial locus-1: an			
RT	embryonic endothelial cell protein that binds the alphabeta3 integrin			
RT	receptor".			
RT	Genes Dev. 12:21-33(1998).			
RL	-1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH			
CC	INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS			
CC	FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN			
CC	REGULATION OF VASCULAR MORPHOGENESIS OF REMODELING IN EMBRYONIC			
CC	DEVELOPMENT.			
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.			
CC	-1- ALTERNATIVE PRODUCTS: TWO ISOFORMS: LONG ISOFORM (SHOWN HERE) AND			
CC	SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.			
DR	EMBL: U70312; AAC02648.1; -;			
DR	EMBL: U70313; AAC02649.1; -;			
DR	HSSP: P00740; IIXA			
DR	INTERPRO: IPR000152; -;			
DR	INTERPRO: IPR000421; -;			
DR	INTERPRO: IPR000561; -;			
DR	INTERPRO: IPR000742; -;			
DR	INTERPRO: IPR001881; -;			

DR PFAM; PF00008; EGF; 3.
 DR PFAM; PF00754; F5_F8_Type C; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 DR EGF-like domain; Alternative splicing; Signal; Developmental protein;
 KW Cell adhesion; Repeat; Vascularization.
 FT SIGNAL 1 16
 FT CHAIN 17 480
 FT DOMAIN 26 59
 FT DOMAIN 78 116
 FT DOMAIN 123 154
 FT DOMAIN 161 311
 FT DOMAIN 322 473
 FT SITE 96 98
 FT DISULFID 26 37
 FT DISULFID 31 48
 FT DISULFID 50 59
 FT DISULFID 78 89
 FT DISULFID 83 105
 FT DISULFID 107 116
 FT DISULFID 158 334
 FT DISULFID 301 305
 FT DISULFID 319 476
 FT VARSPLIC 66 66
 FT VARSPLIC 67 76
 SQ SEQUENCE 480 AA; 53765 MW; F7171E3A309FD48 CRC64;

Query Match 98.4%; Score 2633.5; DB 4; Length 480;
 Best Local Similarity 99.0%; Pred. No. 2.9e-228;
 Matches 476; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 MKRSVAVMLVGLVGLVPGVPGKIDDPNCPNCGICLPGGLAVSGSCCPGFDPNCS 60
 DB 1 MKRSVAVMLVGLVGLVPGVPGKIDDPNCPNCGICLPGGLAVSGSCCPGFDPNCS 60
 QY 61 SYVEVASDEEPTSAQPCPTPNCHNGTCEISEAYRGDTFFIGYVCKPGRFGICQHNH 120
 DB 61 SYVEVASDEEPTSAQPCPTPNCHNGTCEISEAYRGDTFFIGYVCKPGRFGICQHNH 120
 QY 121 NECEVEPCNNGICTDLVANYSCCEPGFERMGNCOYKCSGPLIGEGITISNOQITASSTH 180
 DB 121 NECEVEPCNNGICTDLVANYSCCEPGFERMGNCOYKCSGPLIGEGITISNOQITASSTH 180
 QY 181 RALFGLQKVPYARLNKKGLINAMTAENDRKMIQINQKRRVGVITQGAKRIGS 240
 DB 181 RALFGLQKVPYARLNKKGLINAMTAENDRKMIQINQKRRVGVITQGAKRIGS 240
 QY 241 PEYIKFYKIAVNDGKTAMVYKVGSTNEDMVRGNDINNTPFANSFTPIKAQYRLYPQ 300
 DB 241 PEYIKFYKIAVNDGKTAMVYKVGSTNEDMVRGNDINNTPFANSFTPIKAQYRLYPQ 300
 QY 301 VCRHCHTLMELGELSGCSPEPLMKSGHIDYDITASSIFRTLNMDFTWEPKRALD 360
 DB 301 VCRHCHTLMELGELSGCSPEPLMKSGHIDYDITASSIFRTLNMDFTWEPKRALD 360
 QY 361 KQGVNANTSGHNDOSOMLOVDLVPRTKYTGITTOGADFGHVOGVGYKLAISDGEHW 420
 DB 361 KQGVNANTSGHNDOSOMLOVDLVPRTKYTGITTOGADFGHVOGVGYKLAISDGEHW 420
 QY 421 TVYQDEKORKDKVFOGNFNDTHRRKNVIDPPIYARHRIILPMSWYGRITLASELGTCE 480
 DB 421 TVYQDEKORKDKVFOGNFNDTHRRKNVIDPPIYARHRIILPMSWYGRITLASELGTCE 480
 QY 481 E 481
 DB 480 E 480

RESULT 2
 ID 035474
 AC 035474; PRELIMINARY; PRT; 480 AA.
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DR 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 GN INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.
 OS EDL13 OR DEL1.
 OC Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
 RC TISSUE-EMBRYO;
 RX MEDLINE=98083109; PubMed=9420328;
 RA Hidaï C., Zupancic T., Penta K., Mihail A., Kawana M.,
 RA Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
 RA Auerbach R., Hogan B.L.M., Snodgrass R., Quettermous T.;
 RT "Cloning and characterization of developmental endothelial locus-1: an
 embryonic endothelial cell protein that binds the alphavbeta3 Integrin
 receptor";
 RL Genes Dev. 12:21-33(1998).
 CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
 INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
 FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
 REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC
 DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND
 SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
 CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP
 OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
 NEURONS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. AFTER DAY
 15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME
 OF BIRTH.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC EMBL; AF031524; AAB86585.1; -;
 DR EMBL; AF031525; AAB86586.1; -;
 DR HSSP; P00740; IEDM.
 DR MGD; MGI:1329025; Edl13.
 DR INTERPRO: IPR000152; -;
 DR INTERPRO: IPR000421; -;
 DR INTERPRO: IPR000561; -;
 DR INTERPRO: IPR000742; -;
 DR INTERPRO: IPR001881; -;
 DR PFAM; PF00008; EGF; 3.
 DR PFAM; PF00754; F5_F8_Type C; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_3.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS01285; FAS8C_1; 2.
 DR PROSITE; PS01286; FAS8C_2; 2.
 KW EGF-like domain; Alternative splicing; Signal; Developmental protein;
 KW Cell adhesion; Repeat; Vascularization.
 FT SIGNAL 1 16
 FT CHAIN 17 480
 FT DOMAIN 26 59
 FT DOMAIN 78 116
 FT DOMAIN 123 154
 FT DOMAIN 161 311
 FT DOMAIN 322 473
 FT SITE 96 98
 FT DISULFID 26 37
 FT DISULFID 31 48
 FT DISULFID 50 59
 FT DISULFID 78 89
 FT DISULFID 83 105

FT DISULFID 107 116 BY SIMILARITY.
 FT DISULFID 158 314 BY SIMILARITY.
 FT DISULFID 301 305 BY SIMILARITY.
 FT DISULFID 319 476 BY SIMILARITY.
 FT VARSPLIC 218 221 INLO -> VTWG (IN SHORT ISOFORM).
 FT VARSPLIC 222 480 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 480 AA: 53740 MW: 4CD91EF9261714D CRC64;

Query Match 94.6%; Score 2529.5; DB 11; Length 480;
 Best Local Similarity 94.4%; Pred. No. 6,4e-219;
 Matches 454; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

OY 1 MKRSVAVLVILVGLSGVPGFGKDIQDPNPGENGICLPGIANGSFSCGCEPDEPTPNC 60
 DB 1 MHLVLAAMLVGLSLGVPGFGKDIQDPNPGENGICLPGIANGSFSCGCEPDEPTPNC 60
 OY 61 SVEVVASDEEPTSAQCTPNPCHNGTCEISEAVRGDTFIVGVCKCPRGFNGIHCOHNI 120
 DB 61 SVEVVASDEEPTSAQCTPNPCHNGTCEISEAVRGDTFIVGVCKCPRGFNGIHCOHNI 120
 OY 121 NECEVPCNKNGICTDLVANTSCGCEPFMGRCNQYKCSPLGIEGIIISNOGITASSSTH 180
 DB 121 NECEVPCNKNGICTDLVANTSCGCEPFMGRCNQYKCSPLGIEGIIISNOGITASSSTH 180
 OY 181 RALFGLQKTPYIARLNKGLINANTAAENDRWKRWIOLNQRKMYTGYITGAKRIGS 240
 DB 181 RALFGLQKTPYIARLNKGLINANTAAENDRWKRWIOLNQRKMYTGYITGAKRIGS 240
 OY 241 PEYIKFKIAYSNQDKTAMVYKCTNEDMVFGRNIDNTPYANSFTPIKAQYVRLYPO 300
 DB 240 PEYIKFKIAYSNQDKTAMVYKCTNEDMVFGRNIDNTPYANSFTPIKAQYVRLYPO 300
 OY 301 VCCRHCITLRELLGCLSCSEPLGKSGHIDQVITASSIFRILNMDMTWEPKARLD 360
 DB 301 VCCRHCITLRELLGCLSCSEPLGKSGHIDQVITASSIFRILNMDMTWEPKARLD 360
 OY 360 KQGVNAWMTSGHNDQSQMLQVLDLVPYVYTGITITGAKDFGHVQVSSYKILAYSNDGEHW 420
 DB 360 KQGVNAWMTSGHNDQSQMLQVLDLVPYVYTGITITGAKDFGHVQVSSYKILAYSNDGEHW 420
 OY 421 TVYODEKORBDKVFQGNFNDTHRKNVIDPPIYARHITILFMSWYGRITLASLGLCTE 480
 DB 420 TVYODEKORBDKVFQGNFNDTHRKNVIDPPIYARHITILFMSWYGRITLASLGLCTE 480
 OY 481 E 481
 DB 480 E 480

RESULT 3

O9RLX9

ID O9RLX9 PRELIMINARY: PRT: 463 AA.

AC 09RLX9;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 LONG FORM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=MAMMARY GLAND;
 RX Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
 RT "Lactation-dependent expression of an mRNA splice variant with an exon
 for a multiply O-glycosylated domain of mouse milk fat globule
 glycoprotein MFG-E8".
 RL Biochem Biophys. Res Commun. 254:522-528(1999).
 DR EMBL: AB021130; BAA35180.1; -;
 DR HSSP: P00740; IIXA.
 DR INTERPRO: IPR000421; -;

DR INTERPRO: IPR000561; -;
 DR INTERPRO: IPR001092; -;
 DR INTERPRO: IPR001438; -;
 DR PFAM: PF00008; EGF_2.
 DR PFAM: PF00754; F5_P8_Type_C_2.
 DR PRINTS: PR00010; EGFBLDOD.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 SQ SEQUENCE 463 AA: 51269 MW: D719D2BE090E6427 CRC64;

Query Match 44.0%; Score 1177; DB 11; Length 463;
 Best Local Similarity 49.1%; Pred. No. 1.7e-97;
 Matches 225; Conservative 72; Mismatches 139; Indels 22; Gaps 8;

OY 23 GDICDPNPGENGICLPGIANGSFSCGCEPDEPTPNCSSVEVASDEEPTSAQCTPNP 82
 DB 25 GDCEDSLCLNGGTCILTG-QDNIDYICLPDEGFTGLVONE-----TERGPCSPNP 72
 OY 83 CHNGGTCETS-EAVRGDTFIVGVCKCPRGFNGIHCOHINECEVEPCNKNGICTDLVANY 141
 DB 73 CYNDACGLVLTQGRDIFTEYICQCPVGYSGIHCFETNYNLD--GEYMTAVPWT 129
 OY 142 SCECPGEF--MGRNCOYKCSPLGIEGIIISNOGITASSSTHRALEFGIQRKPYIARLNK 199
 DB 142 SCECPGEF--MGRNCOYKCSPLGIEGIIISNOGITASSSTHRALEFGIQRKPYIARLNK 199
 OY 190 GIVNANTAAENDRWKRWIOLNQRKMYTGYITGAKRIGSPYIKFKIAYSNQDKTWA 259
 DB 190 GIVNANTAAENDRWKRWIOLNQRKMYTGYITGAKRIGSPYIKFKIAYSNQDKTWA 259
 OY 260 MYVKGNGEDMVFGRNIDNTPYANSFTPIKAQYVRLYPOVCRHCTLRELLGCLSG 319
 DB 260 MYVKGNGEDMVFGRNIDNTPYANSFTPIKAQYVRLYPOVCRHCTLRELLGCLSG 319
 OY 249 FIDPE-SGGKEPELGLNDNLSKVNMFLEAOTIRLTPVSCRGCTLRELLGCLSG 307
 DB 249 FIDPE-SGGKEPELGLNDNLSKVNMFLEAOTIRLTPVSCRGCTLRELLGCLSG 307
 OY 320 CSEPLGKSGHIDQVITASSIFRILNMDMTWEPKARLDKQGVNAWMTSGHNDQSQML 379
 DB 320 CSEPLGKSGHIDQVITASSIFRILNMDMTWEPKARLDKQGVNAWMTSGHNDQSQML 379
 OY 308 CSEPLGKSNKTITPDSQMSASSYKTNWLRAGWYPHGLRLDNOCKINAWTAQNSAKEWL 367
 DB 308 CSEPLGKSNKTITPDSQMSASSYKTNWLRAGWYPHGLRLDNOCKINAWTAQNSAKEWL 367
 OY 380 QVLDLVPYVYTGITITGAKDFGHVQVSSYKILAYSNDGEHVTYVQDKORBDKVFQGNP 439
 DB 380 QVLDLVPYVYTGITITGAKDFGHVQVSSYKILAYSNDGEHVTYVQDKORBDKVFQGNP 439
 OY 426 NNSHKNIPEKPMARYVPLVSWHNRITLRELLGC 463
 DB 426 NNSHKNIPEKPMARYVPLVSWHNRITLRELLGC 463

RESULT 4

O9WTS3

ID O9WTS3 PRELIMINARY: PRT: 426 AA.

AC 09WTS3;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 S.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=MAMMARY GLAND;
 RX MEDLINE=99120894; PubMed=9920772;
 RX Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
 RT "Lactation-dependent expression of an mRNA splice variant with an exon
 for a multiply O-glycosylated domain of mouse milk fat globule
 glycoprotein MFG-E8".
 RL Biochem Biophys. Res Commun. 254:522-528(1999).
 DR EMBL: AB025280; BAA76386.1; -;
 DR HSSP: P00740; IIXA.

DR INTERPRO: IPR000421; -
DR INTERPRO: IPR001117; -
DR PFAM: PF00394; Cu-oxidase; 3.
DR PFAM: PF00754; F5_F8_type_C; 2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
SQ SEQUENCE 2183 AA; 247228 MW; BFO8AA723F60317 CRC64;

Query Match 25.5%; Score 682.5; DB 11; Length 2183;
Best Local Similarity 43.2%; Pred. No. 3.5e-52;
Matches 140; Conservative 51; Mismatches 122; Indels 11; Gaps 3;

QY 157 KCSGPLGIEGIIISNOQITASSSTRHALFGLQKWPYARLNKKGLINAMT--AAENDRW 213
DB 1865 ECKMMGSLGVYISDSQIKASEY-----LTWEPRLARLNAGSYNAMSIEKTLDEPI 1918
QY 214 KMWIQINLQKKRVGVITGGAKRIGSEPIYKFIKAIANSNDCKTAMKVKVGTNDWFR 273
DB 1919 KPMIOVDMQKEVYVGIOTGAKHYLKSCTTEFOVAASSDQTNMOIFRGSKGKSVMTFT 1978
QY 274 GNIDNTPYANSEFTPIKQAYRLYPQVCRHCTLMELLGCELSGSEPLGMSGHIQD 333
DB 1979 GNSDGTIKENLDPIYARIYRIHPTKSYNRPTRLRLELQGEVNGCSTPLGLEDRIQD 2038
QY 334 YQITASSIFRTLMDMFTWEPKARLDKQKYNAMTSGHNDOSQWLQVDLVPYVTGII 393
DB 2039 KQITASSFKKSWMGDY--WEPSTARLNAGRYNAWAKANNKQWLOVDLLEIKRYTAVI 2096
QY 394 TGCAKFGHVGQFVSGYKLYANSNGEHWTYQDEKQKDKVFOGNEPDNRKKNVIDPPI 453
DB 2097 TGCKSLSEMYKSYISQISDQGVAKKRYRQKSSMVDKIFEGNSNTKHKHNNFNPPII 2156
QY 454 ARHRIIPMSWYGRITLASLGLC 477
DB 2157 SRFIRIIPKTMQSIALLRLFLFC 2180

RESULT 7
ID 043737 PRELIMINARY; PRT; 2224 AA.
AC 043737;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE FACTOR V.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z99572; CAB16748.1; -
DR HSSP: P00450; 1CRG.
DR INTERPRO: IPR000421; -
DR INTERPRO: IPR001117; -
DR PFAM: PF00394; Cu-oxidase; 3.
DR PFAM: PF00754; F5_F8_type_C; 2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
KW Blood coagulation.
SQ SEQUENCE 2224 AA; 251673 MW; 1013108D49690EAB CRC64;

Query Match 25.1%; Score 670.5; DB 4; Length 2224;
Best Local Similarity 41.8%; Pred. No. 4.3e-51;
Matches 141; Conservative 54; Mismatches 127; Indels 15; Gaps 4;
QY 144 ECPGFEMGRNCOYKCSGPLGIEGIIISNOQITASSSTRHALFGLQKWPYARLNKKGLIN 203

DB 1897 QTEPFLMDRCRM-----PMLSTGIISDSQIKASEY-----LGWEPRLARLNNGSYN 1946
QY 204 AMTA---AENDRWKRWIQINLQKKRVGVITGGAKRIGSEPIYKFIKAIANSNDCKTAM 260
DB 1947 AMSVEKLAFAFASKPMIOVDMQKEVITGOTGAKHYLKSCTTEFOVAASSDQTNMOI 2006
QY 261 YKVGKTEDEWVFPNGINDNTPYANSEFTPIKQAYRLYPQVCRHCTLMELLGCELSG 320
DB 2007 FKGNSTRNVAMFNGNSASTIKENQDFPIYARIYRIHPTKSYNRPTRLRLELQGEVNGC 2066
QY 321 SEPLGMSGHIQDQYQITASSIFRTLMDMFTWEPKARLDKQKYNAMTSGHNDOSQWLQ 380
DB 2067 STPLGMSNGRIENKQITASSFKKSWMGDY--WEPSTARLNAGRYNAWAKANNKQWLE 2124
QY 381 VDLVLPKRVGVITGGAKRIGHVGQFVSGYKLYANSNGEHWTYQDEKQKDKVFOGNEPD 440
DB 2125 IDLTKIKKRIITAIITGCKSLSEMYKSYITIHSEQGEVWKPYRLKSSMVDKIFEGNTI 2184
QY 441 DTHRKNVIDPPIYARHRIIPMSWYGRITLASLGLC 477
DB 2185 KGHYKNEFPNPPIISRFIRIIPKTMQSIALLRLFLFC 2221

RESULT 8
ID 062730 PRELIMINARY; PRT; 2343 AA.
AC 062730;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE FACTOR VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=KIDNEY, SPLEEN;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF049489; AAC05384.1; -
DR HSSP: P00451; 1CRG.
DR INTERPRO: IPR000421; -
DR INTERPRO: IPR001117; -
DR PFAM: PF00394; Cu-oxidase; 3.
DR PFAM: PF00754; F5_F8_type_C; 2.
DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADAADD99 CRC64;

Query Match 24.2%; Score 648; DB 6; Length 2343;
Best Local Similarity 42.6%; Pred. No. 4.8e-49;
Matches 139; Conservative 53; Mismatches 118; Indels 16; Gaps 5;
QY 157 KCSGPLGIEGIIISNOQITASSSTRHALFGLQKWPYARLNKKGLINAMTAAENDRWKRW 216
DB 2031 KCOTPLGMSGHIIRDFQITASSGQY-----GOWAPRLARLHYSGSINAMSTKDP--FSW 2081
QY 217 IQINLQKKRVGVITGGAKRIGSEPIYKFIKAIANSNDCKTAMKVKVGTNDWFRGNI 276
DB 2082 IKVDLAPMTHIGIMTQAROKFSSLYVSOFITMSLDENKWHSTRGNSTGTLMVFFGNV 2141
QY 277 DNNTPYANSEFTPIKQAYRLYPQVCRHCTLMELLGCELSGSEPLGMSGHIQDQYQI 336
DB 2142 DSSGIRKHNIFNPPIIAQYRIHPTKSYNRPTRLRLELQGEVNGCSPMGKESKALSDAQI 2201
QY 337 TASSIFRTLMDMFTWEPKARLDKQKYNAMTSGHNDOSQWLQVDLVPYVTGIIITG 396
DB 2202 TASSYLSM---LATWSPQARLHLOGRTNAMPQANNKREWLQVDFRRTMKVTVGTITG 2258

Db 344 ISKETKKKKYKTYRVADISSNGEDM--ITLKEGNKALIFOGNTNPTDVFVGFPKPLITR 401
 QY 294 YVRLYPQVCRHCHTLRMLLJCELSS--CSEPLGKMSGHQDVOITASSIFRLNMF 351
 Db 402 FVRIKPSWETGISMREFVYGCKITDYPGSGMIGVSGSLSDSQTAS-----NOGRN 455
 QY 352 WEPRKARL--DKQKQVAMT---SGHNDOSQMLQVLDLVPTRKYGITGQAKDFGHQV 406
 Db 456 WMPENIRLVYVSRG---WALPSPHYINEMQLQVLDGDEKIVGVIIOGKHKREKVF 511
 QY 407 GSYLANSNDEGHWTVYODEKQKDKVYQGNFNDTHRKAVIDPEIYARHRLP-WSWY 465
 Db 512 RKFRIASNNQSDMKMIMDSKRAKSFEGNNNYDPELRAF--TPISTREIRIYPERATH 570
 QY 466 GRITLASELCTEE 480
 Db 571 SGLGLRMLLAGEVE 585

RESULT 14
 090W01
 ID 090W01 PRELIMINARY; PRT; 192 AA.
 AC 090W01;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE MILK FAT GLOBULE GLYCOPROTEIN (FRAGMENT).
 GN MFG-E8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oshima K., Aoki N., Negi M., Kishii M., Kitajima K., Matsuda T.;
 RT "Lactation-dependent expression of an mRNA splice variant with an exon
 RT for a multiply O-glycosylated domain of mouse milk fat globule
 RT glycoprotein MFG-E8".
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB021648; BAA36214.1;
 DR EMBL; AB021647; BAA36214.1; JOINED.
 DR INTERPRO: IPR000421;
 DR INTERPRO: IPR000561;
 DR INTERPRO: IPR001092;
 DR PFAM: PF00008; EGF_1.
 DR PFAM: PF00754; F5_F8_Type_C; 1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS01285; FAF5C_1; 1.
 FT NON_TER 1 1
 FT NON_TER 192 192
 SQ SEQUENCE 192 AA; 21134 MW; EEL10C03A89689F4 CRC64;

Query Match 16.0%; Score 428.5; DB 11; Length 192;
 Best Local Similarity 45.4%; Pred. No. 8.7e-31;
 Matches 83; Conservative 33; Mismatches 60; Indels 7; Gaps 4;

QY 77 PCFNPCHNGSTCEIS-EAYRGDPFTGYVCKCPRGFNHICOHNINECEVEPCKNGGICT 135
 Db 1 PCSNPFCYNKAKCLVITDTRGDIFFETIQCQPVGYSGHCEETNYND--GEYMF 57
 QY 136 DLVANSCECPGEF--MGRNCOYKCSGPLGIEGGIISNOITASSTHRALFLQKWPY 193
 Db 58 TAVVNTAVPPAPPLPDSNNLASRCSQTQGMEGALADSQISSASYMGFMGQRMGP 117
 QY 194 ARLNKKLIAMTAENDRKRWIQINLQKRMVYTGVIITGAKRISPEYIKFYKIAVSN 253
 Db 118 ARLYRGIVAMTASNYDS-KPMIOVLLLRKRVSGVMTQASRAGRAEYLTFFKVAAYSL 176
 QY 254 DGK 256
 174

Db 177 DGR 179
 RESULT 15
 ID 061240 PRELIMINARY; PRT; 2352 AA.
 AC 061240;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE HROTCH PROTEIN.
 GN HROTCH.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyuridae; Halocynthia.
 OX NCBI_Taxid=7729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hori S., Saitoh T., Matsumoto M., Makabe K.W., Nishida H.;
 RL Dev. Genes Evol. 207:371-380(1997).
 DR EMBL; AB001327; BAA25571.1;
 DR HSSP; P00740; 1EDM.
 DR INTERPRO: IPR000152;
 DR INTERPRO: IPR000561;
 DR INTERPRO: IPR000742;
 DR INTERPRO: IPR000800;
 DR INTERPRO: IPR001438;
 DR INTERPRO: IPR001881;
 DR INTERPRO: IPR002110;
 DR PFAM: PF00008; EGF_32.
 DR PFAM: PF00023; ank; 6.
 DR PFAM: PF00066; notch; 3.
 DR PRINTS; PR00010; EGFBLD.
 DR PROSITE: PS00010; ASX_HYDROXYL; 18.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_28.
 DR PROSITE: PS01186; EGF_2; 22.
 DR PROSITE: PS01187; EGF_CA; 18.
 KW Glycoprotein; EGF-like domain.
 SQ SEQUENCE 2352 AA; 13DB1C056BBD08D CRC64;

Query Match 12.3%; Score 330; DB 5; Length 2352;
 Best Local Similarity 38.4%; Pred. No. 1.7e-20;
 Matches 63; Conservative 12; Mismatches 47; Indels 42; Gaps 4;
 QY 26 CDPNCPENGLPLGLAVGSFSCPCPDGFTDPNCSVVEVASDEEPTSAAGCTPNPCHN 85
 Db 720 CATQPCQNGTGTSG--INSYNACAPAKYTGNC-----TELSPCVNPCHN 765
 QY 86 GGTCEISEAYRGDTFTGYVCKCPRGFNHICOHNINECEVEPCKNGGICTDLVANYSC 145
 Db 766 GATCOESADY-----LAVYCCQPEGFRGPTCATDINECVNSPCKNGGCTMLVPGYQCTC 820
 QY 146 PGEFMGRNC-----QYKCSGPLGIEG 166
 Db 821 SGGFTGKDCDDIDDCCSSNPCLNGGQCLDDVGSYKCLLPGEFG 864

Search completed: May 23, 2001, 06:20:45
 Job time: 439 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:15:11 ; Search time 75.05 Seconds
(without alignments)
366.363 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513
Perfect score: 2675
Sequence: 1 MKRSVAWLVGLSLGVPOF.....MSWYGRITLASELIGCTEE 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSeq_0401:*
2: /SID6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
3: /SID6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
4: /SID6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
5: /SID6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
6: /SID6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
7: /SID6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
8: /SID6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
9: /SID6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
10: /SID6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
11: /SID6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
12: /SID6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
13: /SID6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
14: /SID6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
15: /SID6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
16: /SID6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
17: /SID6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
18: /SID6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
19: /SID6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
20: /SID6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
21: /SID6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
22: /SID6/gcgdata/geneseq/geneseqp/AA2000.DAT:*
23: /SID6/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2675	100.0	513	20	W94685 Human Del-1 protei
2	2664	99.6	481	18	W10365 Human developmenta
3	2519.5	94.2	480	18	W10364 Mouse developmenta
4	2519.5	94.2	480	20	W94683 Murine Del-1 prote
5	1717	64.2	321	20	W94697 Human milk fat glo
6	1139.5	42.6	221	18	W10366 Murine Del-1 trunc
7	1134.5	42.4	426	21	Y94454 Mouse lactadherin
8	1133.5	42.4	221	20	W94684 Truncated murine D
9	1020.5	38.1	387	16	R77252 HMEG 46 kDa antigen
10	1020.5	38.1	387	21	Y94453 Human lactadherin
11	948	35.4	320	20	W98891 Peptide SEQ ID NO:

12	682.5	25.5	2183	21	B03533 Murine factor V SE
13	675.5	25.3	2224	17	W04254 Human factor V. H
14	670.5	25.1	2224	20	Y49564 Human lipoprotein
15	670	25.0	2319	19	W53485 Murine factor VIII
16	670	25.0	2319	19	W44135 Mus musculus facto
17	670	25.0	2319	19	W31596 Mouse factor VIII
18	670	25.0	2319	22	B50467 Mouse factor VIII.
19	655.5	24.5	2304	21	Y57848 Mouse factor VIII
20	651	24.3	2351	7	P60741 Sequence of human
21	650	24.3	2351	8	P70448 Human factor VIII:
22	650	24.3	2351	16	R78223 Human factor-VIII
23	649	24.3	684	16	R73022 Factor-VIII light
24	649	24.3	684	16	R74091 Factor VIII:c varia
25	649	24.3	1014	8	P71139 Procoagulant-activ
26	649	24.3	1383	18	W33227 Procoagulant-activ
27	649	24.3	1383	18	W33228 Procoagulant-activ
28	649	24.3	1383	18	W33229 Modified factor VI
29	649	24.3	1424	9	P91169 Sequence of 740 Ar
30	649	24.3	1424	10	P91169 Mutant mature huma
31	649	24.3	1424	22	B48842 Modified factor VI
32	649	24.3	1425	9	P80267 B-domain deleted f
33	649	24.3	1438	21	B01262 Factor VIII:SQ.
34	649	24.3	1440	12	R12971 Human factor VIII
35	649	24.3	1457	19	W46246 Human factor VIII
36	649	24.3	1457	19	W44372 Beta-domain delete
37	649	24.3	1457	20	W23414 Human B-domain del
38	649	24.3	1471	18	P80265 Modified factor VI
39	649	24.3	1516	9	P80265 Factor VIII:c varia
40	649	24.3	1661	18	W18670 Factor VIII:c varia
41	649	24.3	2332	8	P71726 Factor VIII:c varia
42	649	24.3	2332	8	P71727 Factor VIII:c varia
43	649	24.3	2332	8	P71728 Factor VIII:c varia
44	649	24.3	2332	8	P71729 Human factor VIII.
45	649	24.3	2332	14	R43257

ALIGNMENTS

RESULT 1	W94685	W94685 standard; Protein: 513 AA.
XX	W94685;	
XX	05-MAY-1999 (first entry)	
DE	Human Del-1 protein.	
XX		
KW	Del-1; developmentally-regulated endothelial cell locus 1; cancer;	
KW	diacidin II; factor VIII-like domain; epidermal growth factor; BGF;	
KW	diabetic retinopathy; rheumatoid arthritis; endometriosis;	
KW	angiogenesis.	
XX		
OS	Homo sapiens.	
PN	US5877281-A.	
XX		
PD	02-MAR-1999.	
XX		
PF	05-JUN-1996; 9605-0659235.	
XX		
PR	05-JUN-1996; 9605-0659235.	
PR	07-JUN-1995; 9505-0480229.	
XX		
PA	(PROG-) PROGENITOR INC.	
PA	(UYVA-) UNIV VANDERBILT.	
XX		
PI	Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;	
XX		
DR	WPI: 1999-189720/16.	
DR	N-PSDB: X18508.	
XX		

PT Del-1 polypeptide sequences - useful for treatment of cancer,
 CC diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Claim 5; Fig 4; 73pp; English.

XX The present sequence is human developmentally-regulated endothelial
 CC cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like
 CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
 CC have an inhibitory effect on angiogenesis (blood vessel growth), this
 CC activity may be useful clinically to prevent neovascularisation of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC integrin, and is an apoptosis inducer.

XX Sequence 513 AA;

Query Match 100.0%; Score 2675; DB 20; Length 513;

Best Local Similarity 100.0%; Pred. No. 2e-163;

Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAWVLVGLSGVPGFGKGDICDPNCCNGICLPGLAYGSPSCCPDFTDPCNS 60
 DB 33 MKRSVAWVLVGLSGVPGFGKGDICDPNCCNGICLPGLAYGSPSCCPDFTDPCNS 92
 QY 61 SVEVASDEEPTSGPCTPNPCNNGTCEISEAVRGDFEIGYVCKCPRGFNGIHCOHNI 120
 DB 93 SVEVASDEEPTSGPCTPNPCNNGTCEISEAVRGDFEIGYVCKCPRGFNGIHCOHNI 152
 QY 121 NCEVEPKNGICITDLYANVSCCEPFGMRNCOYKSCPLIEGIIISNOOITSSRH 180
 DB 153 NCEVEPKNGICITDLYANVSCCEPFGMRNCOYKSCPLIEGIIISNOOITSSRH 212
 QY 181 RALFGLQWRYPYVARLNKKGILINAMTAENDRMKRMIOINLQRMRYTGVITGARRIGS 240
 DB 213 RALFGLQWRYPYVARLNKKGILINAMTAENDRMKRMIOINLQRMRYTGVITGARRIGS 272
 QY 241 PEYIKFYKIAVNSDGTAMKVKVGTNEDWFRGNIDNNPYANSFPPIKAYVLYLQY 300
 DB 273 PEYIKFYKIAVNSDGTAMKVKVGTNEDWFRGNIDNNPYANSFPPIKAYVLYLQY 332
 QY 301 VCRHCTLRMELGCELGCESEPLGMRSGHIDYQITASSIFFTLNMDFTEPRKARLD 360
 DB 333 VCRHCTLRMELGCELGCESEPLGMRSGHIDYQITASSIFFTLNMDFTEPRKARLD 392
 QY 361 KOGKVNAMTSGHNDOSOMLOVDLVPKRVGIIITQGAKGHGFQVSGYKTLAYNSNGEHW 420
 DB 393 KOGKVNAMTSGHNDOSOMLOVDLVPKRVGIIITQGAKGHGFQVSGYKTLAYNSNGEHW 452
 QY 421 TYVODEKORAKVFGQNDNTARRKNVIDPPIYARHRIIRIPWSWYGRITLASELCTEE 480
 DB 453 TYVODEKORAKVFGQNDNTARRKNVIDPPIYARHRIIRIPWSWYGRITLASELCTEE 512
 QY 481 E 481
 DB 513 E 513

RESULT 2
 W10365
 ID W10365 standard; Protein; 481 AA.
 AC W10365;
 XX
 XX 03-MAY-1997 (first entry)
 DE Human developmentally-regulated endothelial cell locus-1 protein.
 XX

KW Del-1; developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnerable; bone formation; diagnosis; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 3..23

XX Protein 24..481

XX Domain 26..59

XX Domain 78..116

XX Domain 123..154

XX Domain 158..315

XX Domain 320..477

XX Domain 481..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

XX Domain 514..513

```

Db      1 mkrsavwllvqlslgvpqfgykdicdpncengicldpqlavgsfscpcpfdpnps 60
QY      61 SYVEVASDEEPTSAQPCPNPCHNGTCEISEAYRGDTFYGVCCKPRGNGIHCOHNI 120
        |||
Db      61 ssvveasdeekpsagpcpnpchngtceiseayrgdtfgyvckcpvgfnglncqnhl 120
QY      121 NECEVEPCKNKGICTDLVANYSCCEPGEFMRNCQYKCSGPIGIEGGIISNOQITASSRH 180
        |||
Db      121 necevepcckngictdlvanysccepgelfmrgncqykcspigiegglisnqglitassrh 180
QY      181 RAIFGLQKWPPYARLNKKGLINAWTAENDRMRWIOINLQKRMKVTVITOGAKRIGS 240
        |||
Db      181 rafifglqkwyppyarlnkkgllnawtaeendrmrwioinlqkrmkvitvigiqgarrijs 240
QY      241 PEYIKFYKTIAYSDGKTWAMYKVKGTNEDMVFNGNIDNNTPYANSTTPPIKAQYVRLYQ 300
        |||
Db      241 peykfyktyiaysdgktwamykvkgtneadmvrfgnidntpyansttprikayvrlpyq 300
QY      301 VCRRHCTLRMELLGCELSGCESEPLGKMSGHIDYQITASSIFRTLMDMFTWPPRKARLD 360
        |||
Db      301 vcrrhctlrmeellgcelsgcseplgkmsghidqyqitassifrtlmdmftwpprkarld 360
QY      361 KOGKVNAMTSGHNDOSQWLOVDLVPRTKVTGIIITOGAKDGHVOFVGSYKLAYSNDGEHW 420
        |||
Db      361 kgkvnamtsgnhndosqwlvdllvpkvtgiltgaakdfghvqfvsykylaysndgehw 420
QY      421 TVYODEKORKDKVFOGNEFNDTHRKNVIDPPYIARHRIILPMSWYGRITLASLCCTEE 480
        |||
Db      421 tvyodekorkdkvfoqnefndthrknavidppiyarhriilpmswygritlaselcgee 480
QY      481 E 481
        |
Db      481 e 481

RESULT      3
W10364
ID      W10364 standard: Protein; 480 AA.
AC      W10364;
XX      03-MAY-1997 (first entry)
DT
DE      XX      Mouse developmentally-regulated endothelial cell locus-1 protein.
XX      Del-1; developmentally-regulated endothelial cell locus-1;
XX      signal transduction; cancer; tumour marker; angiogenesis;
XX      diabetic retinopathy; Rheumatoid arthritis; endometriosi;
XX      cardiac ischaemia; stroke; vascular disease; wound healing;
XX      vulnerary; bone formation; diagnosis; therapy.
OS      Mus sp.
XX
FH      Key      Location/Qualifiers
FT      Peptide 1..23
        /label= Sig_peptide
FT      Protein 24..480
        /label= Mat_protein
FT      Domain 26..59
        /label= EGF-1
        /note= "epidermal growth factor-like domain 1"
FT      Domain 78..116
        /label= EGF-2
        /note= "epidermal growth factor-like domain 2"
FT      Domain 123..154
        /label= EGF-3
        /note= "epidermal growth factor-like domain 3"
FT      Domain 158..314
        /label= Discooidin-1
        /note= "discooidin I/factor VIII-like domain 1"
FT      Domain 319..476
        /label= Discooidin-2
        /note= "discooidin I/factor VIII-like domain 2"
FT

```

```

XX      XX      W09640769-A1.
PN      19-DEC-1996.
PD      05-JUN-1996; 96WO-US09456.
PF      07-JUN-1995; 95US-0480229.
PR      (PROG-) PROGENITOR INC.
PA      (UYVA-) UNIV VANDERBILT.
XX      Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
XX      WPI: 1997-052233/05.
DR      N-PSDB; T47338.
XX      New developmentally regulated endothelial cell locus-1 (del-1) gene
PT      - used to develop prods. for the diagnosis and treatment of cancer
PT      and conditions involving abnormal angiogenesis
PS      Claim 3; Fig 6; 137pp; English.
XX      Murine Del-1 (W10364) is the polypeptide product of the murine
CC      developmentally-regulated endothelial cell locus-1 (del-1) gene
CC      (T47338). It shows 94% amino acid homology to the human Del-1
CC      protein (W10365). Structurally, members of this novel gene family
CC      contain 3 EGF-like domains and 2 discooidin I/factor VIII-like
CC      domains. Del-1 is expressed in endothelial and certain tumour
CC      cells. Its ability to inhibit vascular formation allows its used
CC      as an anti-angiogenic agent. It can be used as a tumour marker,
CC      to identify Del-1 binding partners, and to modulate endothelial
CC      cell growth and blood vessel formation. Recombinant Del-1 can be
CC      produced in transformed host cells utilizing vectors incorporating
CC      del-1 nucleic acids.
XX      Sequence 480 AA:
SQ
Query Match      94.2%; Score 2519.5; DB 18; Length 480;
Best Local Similarity 94.2%; Pred. No. 1,6e-153;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;
QY      1 MKRSVAWLLVIGSLGVPQFGKGDICDPNCPENGICLDGLAVGSFSCCPGFTDPNCS 60
        |||
Db      1 mkrlvaawllvqlslgvpqfgykdicdpncengicldglaavgsfscpcpfdpnps 60
QY      61 SYVEVASDEEPTSAQPCPNPCHNGTCEISEAYRGDTFYGVCCKPRGNGIHCOHNI 120
        |||
Db      61 ssvveasdeekpsagpcpnpchngtceiseayrgdtfgyvckcpvgfnglncqnhl 120
QY      121 NECEVEPCKNKGICTDLVANYSCCEPGEFMRNCQYKCSGPIGIEGGIISNOQITASSRH 180
        |||
Db      121 necevepcckngictdlvanysccepgelfmrgncqykcspigiegglisnqglitassrh 180
QY      181 RAIFGLQKWPPYARLNKKGLINAWTAENDRMRWIOINLQKRMKVTVITOGAKRIGS 240
        |||
Db      181 rafifglqkwyppyarlnkkgllnawtaeendrmrwioinlqkrmkvitvigiqgarrijs 240
QY      241 PEYIKFYKTIAYSDGKTWAMYKVKGTNEDMVFNGNIDNNTPYANSTTPPIKAQYVRLYQ 300
        |||
Db      241 peykfyktyiaysdgktwamykvkgtneadmvrfgnidntpyansttprikayvrlpyq 299
QY      301 VCRRHCTLRMELLGCELSGCESEPLGKMSGHIDYQITASSIFRTLMDMFTWPPRKARLD 360
        |||
Db      301 vcrrhctlrmeellgcelsgcseplgkmsghidqyqitassifrtlmdmftwpprkarld 359
QY      361 KOGKVNAMTSGHNDOSQWLOVDLVPRTKVTGIIITOGAKDGHVOFVGSYKLAYSNDGEHW 420
        |||
Db      361 kgkvnamtsgnhndosqwlvdllvpkvtgiltgaakdfghvqfvsykylaysndgehw 419
QY      421 TVYODEKORKDKVFOGNEFNDTHRKNVIDPPYIARHRIILPMSWYGRITLASLCCTEE 480
        |||

```

Db 420 mvhdeqkrxdkvfgnfdndtrknvldppiyarfillpwsygrltlrseilgcaee 479

QY 481 E 481

Db 480 e 480

RESULT 4

W94683 standard; Protein: 480 AA.

W94683;

05-MAY-1999 (first entry)

Murine Del-1 protein.

Del-1; developmentally-regulated endothelial cell locus 1; cancer; discoidin I; factor VIII-like domain; epidermal growth factor; EGF; diabetic retinopathy; rheumatoid arthritis; endometriosis; angiogenesis.

Mus sp.

US5877281-A.

02-MAR-1999.

05-JUN-1996; 9605-0659235.

05-JUN-1996; 9605-0659235.

07-JUN-1995; 9505-0480229.

(PROG-) PROGENITOR INC. (UTVA-) UNIV VANDERBILT.

Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ; WPI: 1999-189720/16.

N-PSDB: X18506.

Del-1 polypeptide sequences - useful for treatment of cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis

Claim 1; Fig 3; 73pp; English.

The present sequence is murine developmentally-regulated endothelial cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1 proteins have an inhibitory effect on angiogenesis (blood vessel growth), this activity may be useful clinically to prevent neovascularisation of tissues such as tumour nodules and prevention of metastases. The anti-angiogenic activity of Del-1 may be used to treat abnormal conditions that result from angiogenesis, including cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be used to treat cardiac ischaemia, thrombotic stroke, wound healing and peripheral vascular disease. Del-1 is also useful for promoting bone formation. Del-1 binds to alpha V beta 3 Integrin, and is an apoptosis inducer.

Sequence 480 AA;

Query Match 94.2%; Score 2519.5; DB 20; Length 480; Best Local Similarity 94.2%; Pred. No. 1.6e-153; Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

1 MKRSVAWLVLGLSLGVPOFGKGDICDPNCEGNGICLPLAGVSGSCCPDGFDPNCS 60

1 mkhlvaawllvylslgypgfygdcinpcngcngiclsqldadsfsccepegfagpncs 60

61 SVVEVASDEEPTSNAGCTPNPCNGCTCELSKVRDPTTIGVCKCPGFGNKHCHQNI 120

Db 61 svvevasdeekptsagpcipncnglsceliseayrgdtfigyveckprgfnghcqhni 120

QY 121 NECEVEPCKNKGICTDVLVANYSCCEPEFMGRNCOYKCSGPLGEGILISNOQITASTH 180

Db 121 necevepcrngictdvlvanyscepefmgncqykcsgshlgieglisnqglasanh 180

QY 181 RALFGLQKMTPTIARLNKGLINAMTAAENDRKRWIQTINLORRARVGVITQGAKRIGS 240

Db 181 ralfglqkwpyyarlknkgllnawlaendrw-wqdlngtkmrvlgtvgakrligs 239

QY 241 PEYIKFKIAYSNDGKTWAMYKVGKTNEDVFRGNIDNNTPYANSFTPIKAQYRLYPQ 300

Db 240 peyiksykiaysndgktwamykvkgtneemvfrgnvndntpyansftppikayrlypq 299

QY 301 VCRHCTLRMELGCELSGSEPLKMSGHIQDYQTASSIFRTLMQMTWEPKARLD 360

Db 300 lcrhclrmellgcelsgseplgmsghlqdyqltasvfrltnmdftwepkarld 359

QY 361 KQGVNNAWTSQHNDQSOVLVDLVPPTKVTGITQGAQDFGVQFVGSYKLAYSNDGEHW 420

Db 360 kqgvnawtsqhndqsgwlvdlvpkvtgltvgakafghvqfygsklaysndgehw 419

QY 421 TVYQDERQKRDKYFOGNFNDTRKKNVIDPPIYARHRIILPMSWGRITLASELLGCTEE 480

Db 420 mvhdeqkrxdkvfgnfdndtrknvldppiyarfillpwsygrltlrseilgcaee 479

QY 481 E 481

Db 480 e 480

RESULT 5

W94697

W94697 standard; Protein: 321 AA.

W94697;

05-MAY-1999 (first entry)

Human milk fat globule protein MFG-E8.

Del-1; developmentally-regulated endothelial cell locus 1; cancer; discoidin I; factor VIII-like domain; epidermal growth factor; EGF; diabetic retinopathy; rheumatoid arthritis; endometriosis; angiogenesis.

Homo sapiens.

US5877281-A.

02-MAR-1999.

05-JUN-1996; 9605-0659235.

05-JUN-1996; 9605-0659235.

07-JUN-1995; 9505-0480229.

(PROG-) PROGENITOR INC. (UTVA-) UNIV VANDERBILT.

Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ; WPI: 1999-189720/16.

Del-1 polypeptide sequences - useful for treatment of cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis

Example; Column 63-64; 73pp; English.

The present sequence represents human milk fat globule protein MFG-E8, which has homology to the developmentally-regulated endothelial cell locus 1 (Del-1). The Del-1 protein has epidermal growth factor (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1 proteins

[illegible]

QY	322	EPGLMSKCHTODYQIOTASSIFPTLWMDMTTPRRARLIDGKGVNAWTSCHNDSQSLWY	381
Db	273	epgiklntlpdsgmasssytktmrlatqyphlgrlondgklnawtaqnsakewly	332
QY	382	DLLVPKVTGSIITQGAKDFGHVQFSGSYKLAYSNDGEHWIYODEKORDKYRQGFND	441
Db	333	dltqtrvtgltlgdardfghlyvesykvahsdqvgwly--eeqgskvltgndm	390
QY	442	THRKNIWIDPEIYARHRIPLFMSWYGRITLASELLGC	477
Db	391	shkknlfekpfemaryvrvlpvshmrllrllellg	426
RESULT	8		
ID	W94684	standard; Protein: 221 AA.	
AC	W94684		
XX	05-MAY-1999	(first entry)	
XX	Truncated murine Del-1 protein.		
XX	Del-1; developmentally-regulated endothelial cell locus 1; cancer;		
KW	discolidin 1; factor VIII-like domain; epidermal growth factor; EGF;		
KW	diabetic retinopathy; rheumatoid arthritis; endometriosis;		
KW	angiogenesis.		
XX	Mus sp.		
OS	US5877281-A.		
XX	02-MAR-1999.		
XX	05-JUN-1996; 96US-0659235.		
XX	05-JUN-1996; 96US-0659235.		
XX	07-JUN-1995; 95US-0480229.		
XX	(PROG-) PROGENITOR INC.		
XX	(UYVA-) UNIV VANDERBILT.		
XX	Hogan B, Quefermous T, Snodgrass HR, Zupancic TJ;		
XX	WPI, 1999-189720/16.		
XX	N-PSDB; X18507.		
XX	Del-1 polypeptide sequences - useful for treatment of cancer,		
XX	diabetic retinopathy, rheumatoid arthritis and endometriosis		
XX	Claim 3; Column 71-72; 73pp; English.		
XX	The present sequence is truncated murine developmentally-regulated		
XX	endothelial cell locus 1 (Del-1). The protein has epidermal growth factor		
XX	(EGF) like domains and discolidin 1/factor VIII-like domains. The Del-1		
XX	proteins have an inhibitory effect on angiogenesis (blood vessel growth),		
XX	this activity may be useful clinically to prevent neovascularisation of		
XX	tissues such as tumour nodules and prevention of metastases. The anti-		
XX	angiogenic activity of Del-1 may be used to treat abnormal conditions		
XX	that result from angiogenesis, including cancer, diabetic retinopathy,		
XX	rheumatoid arthritis and endometriosis. Since Del-1 promotes		
XX	angiogenesis it can be used to treat cardiac ischaemia, thrombotic		
XX	stroke, wound healing and peripheral vascular disease. Del-1 is also		
XX	useful for promoting bone formation. Del-1 binds to alpha V beta 3		
XX	integrin, and is an apoptosis inducer.		
XX	Sequence 221 AA.		
XX	Query Match 42.4%; Score 1133.5; DB 20; Length 221;		
XX	Best Local Similarity 90.0%; Pred. NO. 2.2e-65;		
XX	Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1		

```

QY 1 MKRSVAWLLVGLSLGVQFGKGDICDPNCENGTICLRLANGLSPSCCEPDPFTDNC 60
DB 1 MKRLVAWLLVGLSLGVQFGKGDICDPNCENGTICLRLANGLSPSCCEPDPFTDNC 60
QY 61 SVVEVASDEEPTASGPCTPNPCNNGTCTEISEAYRGDTFIVGCKPGRFGNCIQHNI 120
DB 61 SVVEVASDEEPTASGPCTPNPCNNGTCTEISEAYRGDTFIVGCKPGRFGNCIQHNI 120
QY 121 NECEVPPCKNGGICTDLVANYSCECPGPFPMGRNOCYKSCGSLGIEGGIISNQOITASTH 180
DB 121 NECEVPPCKNGGICTDLVANYSCECPGPFPMGRNOCYKSCGSLGIEGGIISNQOITASTH 180
QY 181 RALFGLOKWPYRRLANKKGLINMTAENDRKRMQIQL 221
DB 181 RALFGLOKWPYRRLANKKGLINMTAENDRKRMQIQL 221
QY 181 RALFGLOKWPYRRLANKKGLINMTAENDRKRMQIQL 221
DB 181 RALFGLOKWPYRRLANKKGLINMTAENDRKRMQIQL 221

RESULT 9
R77252 standard; Protein: 387 AA.
XX
AC R77252;
XX
DT 21-NOV-1995 (first entry)
XX
DE HMEG 46 kDa antigen.
XX
KW HMEG: human milk fat globule; antigen; immunogen; vaccine; virulence;
KW epithelium; tumor; breast cancer; monoclonal antibody; Mab.
XX
OS Homo sapiens.
XX
PN W09515171-A.
XX
PD 08-JUN-1995.
XX
PE 05-DEC-1994; 94MO-USJ3967.
XX
PR 03-DEC-1993; 93US-0162402.
XX
PA (CANC-) CANCER RES FUND CONTRA COSTA.
XX
PI Ceriani RL, Larocca DJ, Peterson JA;
XX
DR WPI: 1995-215151/28.
XX
N-PSDB; Q91198.
XX
PT 46 kD apparent molecular weight human milk fat globule antigen
PT used in assays to determine the presence of a cancerous tumour of
PT epithelial origin, and in a vaccine against neoplastic tumours
XX
PS Claim 6; Page 46-47; 68pp; English.
XX
CC A complete cDNA sequence for the 46 kDa HMEG antigen, a major
CC component of the apical surface of the normal breast epithelial
CC cell, was obtained by PCR and RACE methods. cDNA clones can be
CC used to prepare MAb's for use in immunotherapy, immunohistochemistry,
CC prognosis, diagnosis, imaging and therapy. Recombinant antigen can
CC be expressed in prokaryotic or (glycosylated) in eukaryotic cells.
XX
SQ Sequence 387 AA;

Query Match 38.1%; Score 1020.5; DB 16; Length 387;
Best Local Similarity 48.8%; Pred. No. 6.4e-58;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;
QY 78 CTNPGHNGTCTEISEAYRGDTFIVGCKPGRFGNCIQHNIINCEHVEPCKNGSICID 136
DB 78 CTNPGHNGTCTEISEAYRGDTFIVGCKPGRFGNCIQHNIINCEHVEPCKNGSICID 136
QY 137 LVANYSCECPGPFPMGRNOCYKSCGSLGIEGGIISNQOITASTHRALEGLQKWPYRRL 196
DB 137 LVANYSCECPGPFPMGRNOCYKSCGSLGIEGGIISNQOITASTHRALEGLQKWPYRRL 196

```

```

DB 66 -----celkveplgmengniansqlaassvrvflgldhwpelari 108
QY 197 NKGILINMTAENDRKRMQIQLINLQKMRVTGVTGCAKRIISPEYIKFYKAIYSDGK 256
DB 109 ntagmynawtrpsndd-npwlqynllrlrmwlyvltgqaslaahaylkaikvayslngn 167
QY 257 TW-AMYKGNEDMPEFGNIDNNTPEYANSFPPIKAQYVRLYPOVCRRCITLAMELIGC 315
DB 168 efdldhvnkshkefy--gwmknayhvnllfetypeagvlylptschactlrlfelligc 225
QY 316 ELSCGSEPLGKSGHIDQVQITASSIFETRLNMDMTPEBRKARILQKQKVNAMTSGHINQ 375
DB 226 elngcamlpikgmsidldkqtlasaykclwylhltswmryarlldqgnfnawagsygn 285
QY 376 SQWLQVLDLVETKYTGITITGAKDFGHVQFVGSYKLAYSNDGEHWYVYODEKORKQKVPQ 435
DB 286 dqwlqvdldgsakevtylltgarnlgsyqfvasykvayandsanwteyqdptryskllp 345
QY 436 GNFDNTHRKNVLDIPYARHRIILPMSWYGRITLASELLGC 477
DB 346 gnwdnshkklfetyplaryrlpvnwvhnrlalrlelligc 387

RESULT 10
Y94453
ID Y94453 standard; Protein: 387 AA.
XX
AC Y94453;
XX
DT 11-SEP-2000 (first entry)
XX
DE Human lactadherin protein.
XX
KW Human; lactadherin; MGF-E8; anti-tumour; immune response;
KW exosome; dendritic cell.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Protein /label= Secretion_signal 24..387
FT /label= Lactadherin 46..48
FT Binding-site /label= Integrin_binding-site
XX
PN EPI004664-A1.
XX
PD 31-MAY-2000.
XX
PE 24-NOV-1998; 98EP-0402925.
XX
PR 24-NOV-1998; 98EP-0402925.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (CURT ) INST CURIE.
XX
DR WPI: 2000-352597/31.
DR N-PSDB: A27140.
XX
PT Chimeric isolated (human) lactadherin polypeptide that functions as an
PT adaptor of cross-priming to eliminate pathogenic antigens, e.g. in
PT tumors -
XX
PS Example 3; Page 12; 20pp; English.
XX
CC Lactadherin protein was found in exosomes produced by dendritic cells.
CC The protein is involved in the phagocytosis of particulate antigens by
CC dendritic cells. Exosomes produced by dendritic cells exposed to
CC tumour antigens induce potent immune responses. Lactadherin or variants
CC of it may be used in the mediation of an immune response. Variants of
CC lactadherin may be used for inhibition and/or stimulation of the
CC cross-priming of antigens and stimulation of the phagocytosis of

```

CC antiens by dendritic cells. Compositions derived from lactadherin can
CC also be used to monitor an immune response, more specifically a CTL
CC (cytotoxic T-lymphocyte) response and also to produce CTLs specific
CC for a selected antigen. The present sequence is the human
CC lactadherin protein.

SQ Sequence 387 AA;

Query Match	38.1%;	Score 1020.5;	DB 21;	Length 387;
Best Local Similarity	48.8%;	Pred. No. 6.4e-58;		
Matches 196;	Conservative 55;	Mismatches 108;	Indels 43;	Gaps 5;

```

QY 78 CTNPONHGTC-ETSEAWRGDTFGVYCKCPFGFHHCOININDECEVECKNGICTD 136
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 27 csknpchngjgleaisgeydvfrpsylctclkyagah----- 65

QY 137 LVANISCECPGEPMRCNQKCSGDLGEGSIISNOOITASSITHRALFELOKWRYPARL 1366
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 66 -----ceckveplgmgenglanlsqiaassayvrlfjlgdwpebarl 1080

QY 197 NKKGLINAWTAAENRMRKRWIOINLOSRMRVTGTYOAGKRGSEYEIKYKIAVNDGK 2566
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 109 nragmwatpssndd-npwqlyvnlrlrmwtygvvgasrlasheylkafvayslng 1676

QY 257 TW-AMKVCVGTEDNAWFGNIDNMTPRVANSTPRIPAKOYRKLRYOVCRRHCTLRKELLG 3151
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 168 efclhdvnhkpkheyl-gnmknaahvnlfrpevaeqyrylprtschactclfiellgc 2251

QY 316 ELSCGSEPLGKMSGHIODYOITASSIFRTLINMDMFEPRKRLDKOGVNAWNTSGHNDQ 3757
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 226 elngcanpigrjknslpdkqltassaykltwjlhfwspysyaldkqgnfawwaagsyn 2855

QY 376 SOWLVQDVLPRKYVGTITOGAKDGFHGVAFSGYKLAISNDGBHMYVYODEKORDKDYF 4353
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 286 dqwlyqvdsgsksevgilttgarnfsgvfaaaykvaaydsanwteyqdprrtsaklfr 3453

QY 436 GNFEDNTHRRKNIWIDPPYARHRIILPMSWGRITTLASELLGC 477
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 346 gwdvnsbhknlfrclprlalyvrlilpvaawhmrilaltellgc 387

```

```

RESULT 11
W98891
ID W98891 standard; peptide: 320 AA.
XX
AC W98891;
XX
DT 05-MAY-1999 (first entry)
EE
DE Peptide SEQ ID NO:20 from US5877281

```

KM Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KM discoidin 1; factor VIII-like domain; epidermal growth factor; EGF;
KM diabetic retinopathy; rheumatoid arthritis; endometriosis;
KM angiogenesis.

OS Unidentified.

PN US5877281-A.

02-MAR-1999
PD

05-JUN-1996; 96US-0659235.

05-JUN-1996: 96075-0659235

PR 07-JUN-1995; 9505-0480229;
XX

PA (PROG-) PROGENITOR INC.

XX

DR WPI; 1999-189720/16.

PT Del-1 polypeptide sequences - useful for treatment of cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis

PS Disclosure; Column 61-62; 73pp; English.

The present invention describes developmentally-regulated endothelial cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) 11 domains and discoidin 1/Factor V/III-like domains. The Del-1 proteins have an inhibitory effect on angiogenesis (blood vessel growth), this activity may be useful clinically to prevent neovascularisation of CC tissues such as tumour nodules and prevention of metastases. The anti-angiogenic activity of Del-1 may be used to treat abnormal conditions that result from angiogenesis, including cancer, diabetic retinopathy, rheumatoid arthritis and endometriosis. Since Del-1 promotes angiogenesis it can be used to treat cardiac ischaemia, thrombotic stroke, wound healing and peripheral vascular disease. Del-1 is also useful for promoting bone formation. Del-1 binds to alpha V beta 3 integrin, and is an apoptosis inducer. The present sequence represents a sequence given in the sequence listing but not mentioned further within the specification.

SQ Sequence 320 AA;

Query Match	35.48;	Score 948;	DB 20;	Length 320;
Best Local Similarity	54.78;	Pred. No. 2.2e-53;		
Matches 175; Conservative	50;	Mismatches 95;	Indels 0;	Gaps 0

OY	158	CSGFLGEGGIIISNOITNASSNRHRLGFLGKMPYARALKSLINAMTAADBNKWKMI	21
OY	1	csqdlgmegalsadqiasasylyumfngldqtwgpeblarlytgcivnanhnaasydxskpwl	60
Db	218	QINLORKMRATGYITNOGAKKRISPEUTYKPKFIAYSDNGKQYAMAYUKKGNEDBYPRGID	277
OY	61	gvnlltkrmvsgvmegsqasragaeaykktkfvayldxrxkfeifqidesgdykelngld	120
Db	278	NNPVPYNSFTPRPKAOYRKLPOVCRHRGCTLRMBELGCELSGSEPRLMSGSHIDVOIT	337
OY	121	nmelkxvnmfnprleeqyrltprvschrgctcltfeilgcelhgscldprldkntlpdsqms	160
Db	338	ASSSTPTLMMDFMTEPRKARLIDKROGVNMTSGHDSQWLOVLRTPRTVYTGIIIOGA	397
OY	181	asssyktmrlafgyrphlgtldngklnawctqmsakewqyldqcgqrqyvtgllbgga	240
Db	398	KDGHGVQFGSGTKFLAYSNDGHEMTVYODEQORXDKPYQSGFNEDNTHRRKNYADPRYLARI	457
OY	241	rdqrlghayvesyvalasaddqywtvxxeegysakvfgnldnmskhknltekrpmayv	300
Db	458	RILPMSWIGRITLASELLGC	477
OY	301	rvlpvshmrlltrtleilg	320

RESULT	12
B03533	
ID	B03533 standard; protein; 2183 AA

DE Murline factor V SEQ ID NO:5.

KW Murine; factor V; FV; activated protein C; APC; anticoagulant;
KW activated protein C resistant factor V; thrombosis; screening;;
KW thrombophilia.

OS Mus sp.

PN US6066778-A

Query Match	Best Local Similarity	Matches	Score	DB	Length	Indels	Gaps
144	25.3%	42.1%	675.5	17	2224	15	4
1897	142	54	126	15	1946	15	4
204	147	54	126	15	1946	15	4
1947	147	54	126	15	1946	15	4
261	147	54	126	15	1946	15	4
2007	147	54	126	15	1946	15	4
321	147	54	126	15	1946	15	4

```

Dn      2067  sdrplamgnkiknkgqitlaasfkkswgdy--weftralnagryrnagakanmkkwyle 2124
Qy      381  vdllyvtrkgtvigtogakdpgvhvofvgsykylasNDGDEHWTYODEKORKDKVFGNFDN 440
Db      2125  idllkktkaltitgckelsesemykayfthlysegyewkpyrlkssmvdikfegntnt 2184
Qy      441  DTHRKNVIDPPYIARHRIILPMSWGRTILASELLGC 477
Db      2185  kghvknfipnplisrflrvipkrtmgsiltlrlelfgc 2221

RESULT  14
ID      Y49564 standard; Protein; 2224 AA.
XX      Y49564;
AC      Y49564;
DT      13-JAN-2000 (first entry)
XX      Human lipoprotein lipase protein sequence.
DE      Human: coding sequence polymorphism; vascular pathology gene;
KW      polymorphic site; phenotype correlation; forensic; paternity testing;
KW      medicine; genetic analysis; vascular disease.
OS      Homo sapiens.
XX      MO9950454-A2.
XX      PN
XX      PD 07-OCT-1999.
XX      PF 26-MAR-1999; 99WO-US06473.
XX      PR 01-APR-1998; 98US-0054272.
XX      (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA      Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;
PI      WPI; 1999-620066/53.
DR      N-PSDB; Z32183.
XX      Determination of polymorphisms in genes, especially those identifying
PT      predisposition to vascular disease
XX      Disclosure; Fig 27; 134pp; English.
XX      PS
XX      CC Z32159 to Z32194 represent reference alleles for specifically claimed
CC      nucleic acid sequences from the present invention which comprise
CC      polymorphic sites as given in a table in the specification, selected
CC      from 92 single nucleotide polymorphisms in which the nucleotide at the
CC      polymorphic site is different from a nucleotide at the same site in a
CC      reference allele. The nucleic acids, and primers and probes, are used to
CC      identify polymorphisms, which may predispose an individual to disease,
CC      especially a vascular disease. They can also be used in phenotype
CC      correlations, forensics, paternity testing, medicine or genetic
CC      analysis. Y49550 to Y49573 represent the proteins which correspond
CC      to some of the reference alleles.
XX      CC
XX      SQ Sequence 2224 AA;

Query Match 25.1%; Score 670.5; DB 20; Length 2224;
Best Local Similarity 41.8%; Pred. No. 8,9e-35;
Matches 141; Conservative 54; Mismatches 127; Indels 15; Gaps 4;

Qy      144  ECGEFGMRNCQYKCGSPGLGIEGGITNSOQITASSTHRALFELQKWYRYARLNKKGLIN 203
Db      1897  qtfllimdrdrdm-----puglstgltisdqikasef-----lyweprlarlngngsyn 1946
Qy      204  AWTAA--AENDRRKKRIQIINDLRKKMNVYGVITOGAKRIGSPYIINFYKIAYSNDOKTAM 260

```

Db 1947 awsvkelaefaskpwiqdmqkewilgtqgahy1kscytefyvayssnqinwi 2006
 QY 261 YVKKGNEEDMVRGNIDNNTFPYANSETPPIKAQYVRLVPOVCRHCTLRMELGCELSGC 320
 Db 2007 fkgnsrnmwmytngnsdastlkengfdppivayrlrslptaynrplrlleqgevcngc 2066
 QY 321 SPSPLGKSGHIDYQITATTSIFRTLNMDFTWEPKARLDKQKVNAMTSGHNDQSOMIQ 380
 Db 2067 stpplmgengkienkqitassfkswgdy--wefirarlmgayrvnawgakannmkqyle 2124
 QY 381 VDLVPTKVTGIIITGACDFGHVQVGSYKLAYSNDGEHWTYQDEKORKDKVFOGNEFDN 440
 Db 2125 idlllkkikitalitgcsksissemwkyrlhysggyewekpyrlkssmvdikfegntnt 2184
 QY 441 DTHRRKVIDPPIYARHRIPLPMSWGRITLASELLGC 477
 Db 2185 kghvknfnpplisrlirvlpkwtngsalrllelgtc 2221
 RESULT 15
 W53485
 ID W53485 standard; Protein: 2319 AA.
 AC W53485;
 DT 10-JUL-1998 (first entry)
 DE Murine factor VIII.
 KW Murine; factor VIII; hybrid; haemophilia; procoagulant; blood; clot.
 OS Mus musculus.
 PN US5744446-A.
 PD 28-APR-1998.
 PF 07-JUN-1995; 95US-0474503.
 PR 07-JUN-1995; 95US-0474503.
 PR 07-APR-1992; 92US-0864004.
 PR 11-MAR-1994; 94US-0212133.
 PR 15-NOV-1994; 94WO-US13200.
 PA (UYEM-) UNITV EMORY.
 PI Lollar JS, Runge MS;
 DR WPI: 1998-271107/24.
 DR N-PSDB; V25812.
 PT Hybrid of human and animal factor VIII - containing porcine and
 PT murine amino acid sequences is useful in the treatment of
 PT haemophilia
 PS Claim 9; Column 73-84; 48pp; English.
 XX The present sequence represents murine factor VIII used in the present
 CC invention. The present invention describes a new procoagulant hybrid
 CC factor VIII (I) comprising human factor VIII and has amino acid
 CC sequences substituted from the group of A2 domain fragments consisting
 CC of amino acids 373-540, 373-508, 445-508, 484-508, 404-508, 489-508
 CC and 484-489 from the human 2332 amino acid A2 domain sequence (II) as
 CC given in the specification (see w53483). The substitution is from
 CC corresponding non-human mammalian factor VIII sequences. Also described
 CC is a method for treating factor VIII deficiency comprising administering
 CC a therapeutically effective dose of (I) in a pharmaceutical carrier.
 CC (I), prepared from reconstitution of purified molecules or recombinant
 CC techniques, is useful in the treatment of haemophiliacs who have factor
 CC VIII deficiencies and whose blood is not normally able to clot after
 CC internal or external bleeding. (I) compared to native human factor VIII
 CC is more stable at physiological conditions and has a higher specific
 CC clotting activity.

XX Sequence 2319 AA;
 SQ
 Query Match 25.0%; Score 670; DB 19; Length 2319;
 Best Local Similarity 46.3%; Pred. No. 1e-34;
 Matches 151; Conservative 40; Mismatches 119; Indels 16; Gaps 5;
 QY 157 KCSPLGIEGGITISNOCITASSSTRALFGLQKWPYARLNKGLINAMTAENDMKRR 216
 Db 2007 qcqplgmasgsirfdqitassghy-----gwaopnalarlhysgslnawstkep---fsw 2057
 QY 217 IOINLORKMRVYTGITGAKRIGSPXYIKFYKLAYSNDGKTMAMKYKGTNEEDMVRGNT 276
 Db 2058 ikvdlapmavhykigtgarqkfslylsqfilmysldgkwlvsyqgnstgltlmvfgnv 2117
 QY 277 DNTFPYANSETPPIKAQYVRLVPOVCRHCTLRMELGCELSGSGSEPLGKSGHIDYQI 336
 Db 2118 dsqgikhnsfnplliarylrlhplhsslsrlmelngcdlnscslplgmeskvisdtq1 2177
 QY 337 TASSIFRTLNMDFTWEPKARLDKQKVNAMTSGHNDQSOMLOVDLVPKVTGIIITG 395
 Db 2178 tassyf---tmfatawpsqarlhlqgrtnawrpyvnpkqwlqvd1qktmkvvtglltq 2233
 QY 396 GAKDFGHVQVGSYKLAYSNDGEHWTYQDEKORKDKVFOGNEFDNTHRRKVIDPPIYAR 455
 Db 2234 gvkelftsmfvkeflissqgqghwt--qillyngkvavfgngdsstpmmsldppllrr 2291
 QY 456 HIRLPMSWGRITLASELLGCTEE 481
 Db 2292 ylrhlpqiwehqlalrllelgtceaq 2317

Search completed: May 23, 2001, 06:15:14
 Job time: 228 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:18:12 ; Search time 179.67 Seconds

(without alignments)
430.629 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513

Perfect score: 2675
Sequence: 1 MKRSYAVWLIVGLSLGVPQF.....MSWYGRITLASELLGCTEEH 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*

2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*

3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*

4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*

6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*

8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*

9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*

10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*

11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*

12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*

13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*

14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*

15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*

16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*

17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*

18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*

19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*

20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*

21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*

22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*

23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2675	100.0	513	8	US-08-480-2298-14
2	2675	100.0	513	10	US-08-659-2358-14
3	2675	100.0	513	16	US-09-237-981-14
4	2646	98.9	481	10	US-08-659-2358-29
5	2519.5	94.2	480	8	US-08-480-2298-10
6	2519.5	94.2	480	10	US-08-659-2358-10
7	2519.5	94.2	480	16	US-09-237-981-10
8	1717	64.2	321	8	US-08-480-2298-21
9	1717	64.2	321	10	US-08-659-2358-21
10	1717	64.2	321	16	US-09-237-981-21

11	1615	52.6	449	23	US-60-207-315-404	Sequence 404, App
12	1406.5	52.6	362	23	US-60-207-315-523	Sequence 523, App
13	1405.5	52.5	311	23	US-60-230-435-1519	Sequence 1519, App
14	1136	43.2	463	19	US-09-582-340-4	Sequence 4, App1
15	1133.5	42.4	221	16	US-09-237-981-29	Sequence 29, App1
16	1039.5	38.9	203	23	US-60-233-644-77	Sequence 77, App1
17	1036.5	38.7	185	23	US-60-230-435-1806	Sequence 1806, App
18	1020.5	38.1	387	19	US-09-582-340-2	Sequence 2, App1
19	948	35.4	320	8	US-08-480-2298-20	Sequence 20, App1
20	948	35.4	320	10	US-08-659-2358-20	Sequence 20, App1
21	948	35.4	320	16	US-09-237-981-20	Sequence 20, App1
22	675.5	25.3	2224	1	PCT-US95-03881A-14	Sequence 14, App1
23	675.5	25.3	2224	8	US-08-410-488-14	Sequence 14, App1
24	670.5	25.1	2224	14	US-09-054-272-38	Sequence 14, App1
25	670.5	25.1	2228	23	US-60-208-020-170	Sequence 170, App
26	670.5	25.1	2228	23	US-60-209-043-221	Sequence 221, App
27	670.5	25.1	2228	23	US-60-213-172-167	Sequence 167, App
28	670.5	25.1	2228	23	US-60-213-173-167	Sequence 167, App
29	670	25.0	2319	1	PCT-US00-13541-6	Sequence 6, App1
30	670	25.0	2319	1	PCT-US01-05076A-28	Sequence 28, App1
31	670	25.0	2319	1	PCT-US97-11155-6	Sequence 6, App1
32	670	25.0	2319	8	PCT-US99-05193-6	Sequence 6, App1
33	670	25.0	2319	8	US-08-475-201-6	Sequence 6, App1
34	670	25.0	2319	17	US-09-315-179-6	Sequence 6, App1
35	670	25.0	2319	19	US-09-523-656-28	Sequence 28, App1
36	655.5	24.5	2304	14	US-09-035-141-4	Sequence 4, App1
37	655.5	24.5	2304	17	US-09-324-867-4	Sequence 4, App1
38	649	24.3	1438	1	PCT-US93-29169-1	Sequence 1, App1
39	649	24.3	1438	16	US-09-209-916-1	Sequence 1, App1
40	649	24.3	1457	12	US-08-669-309-47	Sequence 47, App1
41	649	24.3	1457	14	US-09-001-039A-47	Sequence 47, App1
42	649	24.3	1457	14	US-09-001-039A-47	Sequence 47, App1
43	649	24.3	1457	15	US-09-190-941-47	Sequence 47, App1
44	649	24.3	1471	10	US-08-683-839A-3	Sequence 3, App1
45	649	24.3	1661	16	US-09-243-532-2	Sequence 2, App1

ALIGNMENTS

RESULT 1

US-08-480-2298-14

Sequence 14, Application US/08480229B

GENERAL INFORMATION:

APPLICANT: Quartermours, Thomas

APPLICANT: Hogan, Bridgid

APPLICANT: Snodgrass, H. Ralph

APPLICANT: Zupancic, Thomas J.

TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/480,229B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.

REGISTRATION NUMBER: 28, 462

REFERENCE/DOCKET NUMBER: 8907-026

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-2298-14

Query Match 100.0%; Score 2675; DB 8; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.6e-221;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAWMLVGLSLGVPOGKGDICDPNCPENGICLPGIAGVSFSCPDGFTDNC 60
DB 33 MKRSVAWMLVGLSLGVPOGKGDICDPNCPENGICLPGIAGVSFSCPDGFTDNC 92
QY 61 SVEVASDEEPTPSAGCPTPNPCNNGTCSEISAYRGDTFTIGYCKCPRGNGIHCOHNI 120
DB 93 SVEVASDEEPTPSAGCPTPNPCNNGTCSEISAYRGDTFTIGYCKCPRGNGIHCOHNI 152
QY 121 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPLGIEGIIISMOQTASTH 180
DB 153 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPLGIEGIIISMOQTASTH 212
QY 181 RALFGLQKWPYYARLKKGLINAMTAENDRMKRWIOLNLRKRVGTGITGAKRIGS 240
DB 213 RALFGLQKWPYYARLKKGLINAMTAENDRMKRWIOLNLRKRVGTGITGAKRIGS 272
QY 241 PEYIKFKIAYSNDGKTWAMYKVGKTNEEDMVEFGINDNTPYANSFTPTIAQYVRLYPQ 300
DB 273 PEYIKFKIAYSNDGKTWAMYKVGKTNEEDMVEFGINDNTPYANSFTPTIAQYVRLYPQ 332
QY 301 VCRHCHTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRLLNMDMFTWEPKRALD 360
DB 333 VCRHCHTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRLLNMDMFTWEPKRALD 392
QY 361 KOGVNAWMTSGHNDOSQWLOVDLLVPTKYTGIIITGAKDFGHVOFVSGYKLAYSNDEHW 420
DB 393 KOGVNAWMTSGHNDOSQWLOVDLLVPTKYTGIIITGAKDFGHVOFVSGYKLAYSNDEHW 452
QY 421 TVYODEKORKDKYFGNFDNDRKKNVIDPPIYARHIRILPMSWGRITTLASELLGCTEE 480
DB 453 TVYODEKORKDKYFGNFDNDRKKNVIDPPIYARHIRILPMSWGRITTLASELLGCTEE 512
QY 481 E 481
DB 513 E 513

RESULT 2
US-08-659-2358-14

Sequence 14, Application US/086592358
GENERAL INFORMATION:
APPLICANT: Quartermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENOPTHETAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-2358-14

Query Match 100.0%; Score 2675; DB 10; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.6e-221;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAWMLVGLSLGVPOGKGDICDPNCPENGICLPGIAGVSFSCPDGFTDNC 60
DB 33 MKRSVAWMLVGLSLGVPOGKGDICDPNCPENGICLPGIAGVSFSCPDGFTDNC 92
QY 61 SVEVASDEEPTPSAGCPTPNPCNNGTCSEISAYRGDTFTIGYCKCPRGNGIHCOHNI 120
DB 93 SVEVASDEEPTPSAGCPTPNPCNNGTCSEISAYRGDTFTIGYCKCPRGNGIHCOHNI 152
QY 121 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPLGIEGIIISMOQTASTH 180
DB 153 NECEVEPCCKNGICTDLVANYSCCEPGEPMGRNCOYKCSGPLGIEGIIISMOQTASTH 212
QY 181 RALFGLQKWPYYARLKKGLINAMTAENDRMKRWIOLNLRKRVGTGITGAKRIGS 240
DB 213 RALFGLQKWPYYARLKKGLINAMTAENDRMKRWIOLNLRKRVGTGITGAKRIGS 272
QY 241 PEYIKFKIAYSNDGKTWAMYKVGKTNEEDMVEFGINDNTPYANSFTPTIAQYVRLYPQ 300
DB 273 PEYIKFKIAYSNDGKTWAMYKVGKTNEEDMVEFGINDNTPYANSFTPTIAQYVRLYPQ 332
QY 301 VCRHCHTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRLLNMDMFTWEPKRALD 360
DB 333 VCRHCHTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRLLNMDMFTWEPKRALD 392
QY 361 KOGVNAWMTSGHNDOSQWLOVDLLVPTKYTGIIITGAKDFGHVOFVSGYKLAYSNDEHW 420
DB 393 KOGVNAWMTSGHNDOSQWLOVDLLVPTKYTGIIITGAKDFGHVOFVSGYKLAYSNDEHW 452
QY 421 TVYODEKORKDKYFGNFDNDRKKNVIDPPIYARHIRILPMSWGRITTLASELLGCTEE 480
DB 453 TVYODEKORKDKYFGNFDNDRKKNVIDPPIYARHIRILPMSWGRITTLASELLGCTEE 512
QY 481 E 481
DB 513 E 513

RESULT 3
US-09-237-981-14

Sequence 14, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quartermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph

APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-14

Query Match 100.0%; Score 2675; DB 16; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.6e-221;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCGNGICLPGLAVGSFSCPCPDGFTDPNCS 60
DB 33 MKRSVAVMLVGLSLGVPOFGKGDICDPNCGNGICLPGLAVGSFSCPCPDGFTDPNCS 92

QY 61 SVVEVAVSDEEPTSAAGCPNCPCHNGTCEISEAYRGDTFIGYVCKCPRGFGNGIHCQHN 120
DB 93 SVVEVAVSDEEPTSAAGCPNCPCHNGTCEISEAYRGDTFIGYVCKCPRGFGNGIHCQHN 152

QY 121 NCEVEVPCCKNGICITDLVANYSCCEPGEFMRNCQYKCSGPGITGEGGIIISNOQITASSTH 180
DB 153 NCEVEVPCCKNGICITDLVANYSCCEPGEFMRNCQYKCSGPGITGEGGIIISNOQITASSTH 212

QY 181 RALFGLQKWPYYARLNKKGLINAMTAENDRMRKMIQINLQKRMRYGVITOGAKRIGS 240
DB 213 RALFGLQKWPYYARLNKKGLINAMTAENDRMRKMIQINLQKRMRYGVITOGAKRIGS 272

QY 241 PEYIKFKIYISNDGKTWAMYKVGTTNEDVAFRGNTDNNTPYANSFTPPKKAQYVRLYRQ 300
DB 273 PEYIKFKIYISNDGKTWAMYKVGTTNEDVAFRGNTDNNTPYANSFTPPKKAQYVRLYRQ 332

QY 301 VCRHCHTLRMELLGCELSGSEPLGKMSGHIQDYQITASSIFPTLNKDMFTWEPKARLD 360
DB 333 VCRHCHTLRMELLGCELSGSEPLGKMSGHIQDYQITASSIFPTLNKDMFTWEPKARLD 392

QY 361 KQGVNMTSGHNDQSWLOVDLVPYKVGITITGQAKDGHYQVFCSTYLASNDGEHW 420
DB 393 KQGVNMTSGHNDQSWLOVDLVPYKVGITITGQAKDGHYQVFCSTYLASNDGEHW 452

QY 421 TVYDEKQKDKVQGNEDNDTHRKNVIDPPIYARHIRILPMSWYGRITLASLCTEE 480
DB 453 TVYDEKQKDKVQGNEDNDTHRKNVIDPPIYARHIRILPMSWYGRITLASLCTEE 512

QY 481 E 481
DB 513 E 513

RESULT 4
US-08-659-235B-29
Sequence 29, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-29

Query Match 98.9%; Score 2646; DB 10; Length 481;
Best Local Similarity 99.2%; Pred. No. 4.6e-219;
Matches 477; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCGNGICLPGLAVGSFSCPCPDGFTDPNCS 60
DB 1 MKRSVAVMLVGLSLGVPOFGKGDICDPNCGNGICLPGLAVGSFSCPCPDGFTDPNCS 60

QY 61 SVVEVAVSDEEPTSAAGCPNCPCHNGTCEISEAYRGDTFIGYVCKCPRGFGNGIHCQHN 120
DB 61 SVVEVAVSDEEPTSAAGCPNCPCHNGTCEISEAYRGDTFIGYVCKCPRGFGNGIHCQHN 120

QY 121 NCEVEVPCCKNGICITDLVANYSCCEPGEFMRNCQYKCSGPGITGEGGIIISNOQITASSTH 180
DB 121 NCEVEVPCCKNGICITDLVANYSCCEPGEFMRNCQYKCSGPGITGEGGIIISNOQITASSTH 180

QY 181 RALFGLQKWPYYARLNKKGLINAMTAENDRMRKMIQINLQKRMRYGVITOGAKRIGS 240
DB 181 RALFGLQKWPYYARLNKKGLINAMTAENDRMRKMIQINLQKRMRYGVITOGAKRIGS 240

OY	241	PEYIEFYKIAVSNCKCTAAMTKVKGVTNEDMVEFGINDINNTPYANSPPIKAOVRLPYO	3000
Db	241	PEYIEFYKIAVSNCKCTAAMTKVKGVTNEDMVEFGINDINNTPYANSPPIKAOVRLPYO	3000
OY	301	VCRRHCTRLMELLCGELSGCSEPLGMSGHIQDYQITASSIFTRLNMDMFTWBRKARLD	3600
Db	301	VCRRHCTRLMELLCGELSGCSEPLGMSGHIQDYQITASSIFTRLNMDMFTWBRKARLD	3600
OY	361	KQKKNAMTSGHNOSQWLOYDVLVPTKVGITITQGAKDGHQVQVGSYKLAYSNGEHW	4200
Db	361	KQKKNAMTSGHNOSQWLOYDVLVPTKVGITITQGAKDGHQVQVGSYKLAYSNGEHW	4200
OY	421	TVYODEKQKDVQGNFENDNTHRKNYIDPEIYARHRIILPWSWYGRITLASELCTTEE	4800
Db	421	TVYODEKQKDVQGNFENDNTHRKNYIDPEIYARHRIILPWSWYGRITLASELCTTEE	4800
OY	481	E 481	
Db	481	E 481	

RESULT 5
US-08-480-229B-10
; Sequence 10, Application US/08480229B
GENERAL INFORMATION

```

:
:
: APPLICANT: Quertermous, Thomas
: APPLICANT: Hogan, Bridgid
: APPLICANT: Snodgrass, H. Ralph
: APPLICANT: Zupancic, Thomas J.
:
: TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
:
: TITLE OF INVENTION: CELL LOCUS-1
:
: NUMBER OF SEQUENCES: 28
:
: CORRESPONDENCE ADDRESS:
:

```

```

ADDRESSSEE: Pennine & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Polissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66144 Pennine
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
OS-08-480-229B-10

```

Query Match	94.2%;	Score 2519.5;	DB 8;	Length 480;
Best Local Similarity	94.2%;	Pred. No. 3.5e-208;		
Matches 453; Conservative	9;	Mismatches 18;	Indels 1;	Gaps 1

[illegible]

Qy	61	SVEVAASDEEPTSGPCTPMDCHNGGTCETSEAYRBDTETGYCKCPDRPNGIHOHNT	120
Db	61	SVEVAASDEEPTSGAPCIPNCHNGGTCTSEAYRBDTETGYCKCPDRPNGIHOHNT	120
Qy	121	NECEVEPCKNGSICITDLVANTSCDEOPGFMRNCOYKCSGPLIEGIIISNOOITASSTH	180
Db	121	NECEAEPRNGRGICTDLVANTSCDEOPGFMRNCOYKCSGLIEGIIISNOOITASSNH	180
Qy	181	RLAELOKMYPPYALANKKGLINMTAENDRMKRWQINLQRMKRTGYITOGAKRIGS	240
Db	181	RLAELOKMYPPYALANKKGLINMTAENDRW- WQINLQRMKRTGYITOGAKRIGS	239
Qy	241	PEYIEFYKIAYSNDKTMAMTKVAGTDEDMYFRGININNTPYANSFPPKKAQYVRLPYQ	300
Db	240	PEYIKSYKIAYSNDKTMAMTKVAGTDEMYFRGVNNTNPNYANSFPPKKAQYVRLPYQ	299
Qy	301	VCRRHCTRLMELGCELSGCSGEPJGMSGHIQDOYITASSIFPTLNDMTWEPERRARLD	360
Db	300	ICRRHCTRLMELGCELSGCSGEPJGMSGHIQDOYITASSVFPTLNDMTWEPERRARLD	359
Qy	361	KQGNKNAVNTSGHNOSOMLQYDVLVPRKVGIIITOGAKDGHQVFGYSYLAASNGEHW	420
Db	360	KQGNKNAVNTSGHNOSOMLQYDVLVPTKVGIIITOGAKDGHQVFGYSYLAASNGEHW	419
Qy	421	TVYODEKQKRDVFGQFNDDNTHRKNAVJDPYIYARHRIILPMSMYGRITLASELCTCEE	480
Db	420	MVHODEKQKRDVFGQFNDDNTHRKNAVJDPYIYARFRIILPMSMYGRITLRSLLCAAE	479
Qy	481	E 481	
Db	480	E 480	

RESULT 6
US-08-659-235B-10
Sequence 10

; sequence 10, Application US/08659435B
; GENERAL INFORMATION:

1 APPLICANT: Quartermoss, Thomas
2 APPLICANT: Hogan, Brigid
3 APPLICANT: Snodgrass, H. Ralph
4 APPLICANT: Zupancic, Thomas J.
5 TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
6 TITLE OF INVENTION: CELL LOCUS-1
7 NUMBER OF SEQUENCES: 29
8 CORRESPONDENCE ADDRESS: 9

```

?      ADDRESSEE: Pennie & Edmonds LLP
?      STREET: 1155 Avenue of the Americas
?      CITY: New York
?      STATE: New York
?      COUNTRY: United States
?      ZIP: 10036-2711
?      COMPUTER READABLE FORM:

```

1 MEDIUM TYPE: Floppy disk
 2 COMPUTER: IBM PC compatible
 3 OPERATING SYSTEM: PC-DOS/MS-DOS
 4 SOFTWARE: PatentIn Release #1.0, Version #1.30
 5 CURRENT APPLICATION DATA:
 6 APPLICATION NUMBER: US/08/659,235B
 7 FILING DATE: 05-JUN-1996
 8 CLASSIFICATION: 435

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Poissant, Brian M.
:
: REGISTRATION NUMBER: 28,462
: REFERENCE/DOCKET NUMBER: 8907-03
: TELECOMMUNICATION INFORMATION:
:

```

```

; TELEPHONE: (212) 790-9090
;
; TELEFAX: (212) 869-8864/97
;
; TELEX: 66141 Pennie
;
; INFORMATION FOR SEQ ID NO: 10:
;
; SEQUENCE CHARACTERISTICS:

```

```

;          LENGTH: 480 amino acids
;          TYPE:  amino acid
;          TOPOLOGY:  linear
;

```


MOLECULE TYPE: protein
US-08-659-235B-10

Query Match 94.2%; Score 2519.5; DB 10; Length 480;
Best Local Similarity 94.2%; Pred. No. 3.5e-208;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

```
OY 1 MKRSVAWLLVGLSLGVPOFGKGDICDPNCPENGICLPGIAGVSFSCPCDGFDPNCS 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MKHLVAAMLVLVGLSLGVPOFGKGDICDPNCPENGICLGLADSFSCPCDGFDPNCS 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 SVEEASDEEPTSGPCTPNPCCHNGTCEISEAYRDTFTGYCKCPRGNGIHCOHNI 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 SVEEASDEEPTSGPCTPNPCCHNGTCEISEAYRDTFTGYCKCPRGNGIHCOHNI 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 NECEVEPCRNKGICITDLVANYSCCEPGEFMRNCQKCSGGLGIEGIIISNOQITASSH 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 NECEVEPCRNKGICITDLVANYSCCEPGEFMRNCQKCSGGLGIEGIIISNOQITASSH 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 RALFGLQKWPYYARLNKKGILNMTAENDRMWIDQINQRMRYGVITOGAKRIGS 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 RALFGLQKWPYYARLNKKGILNMTAENDRMWIDQINQRMRYGVITOGAKRIGS 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 241 PEYIKFYKIAVNDGKTAMVKYKGTNEDMVRGNIDNNTPYANSFTPIKAQYVRLYPQ 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 PEYIKFYKIAVNDGKTAMVKYKGTNEDMVRGNIDNNTPYANSFTPIKAQYVRLYPQ 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 301 VCRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 VCRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 300 ICRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 300 ICRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 361 KQKYNAMTSGHNDOSOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHM 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 360 KQKYNAMTSGHNDOSOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHM 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 421 TVYODEKQKDKVFOGNEFDNTHRNKVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 420 TVYODEKQKDKVFOGNEFDNTHRNKVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 481 E 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 480 E 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 7
US-09-237-981-10
Sequence 10, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-10

Query Match 94.2%; Score 2519.5; DB 16; Length 480;
Best Local Similarity 94.2%; Pred. No. 3.5e-208;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

```
OY 1 MKRSVAWLLVGLSLGVPOFGKGDICDPNCPENGICLPGIAGVSFSCPCDGFDPNCS 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MKHLVAAMLVLVGLSLGVPOFGKGDICDPNCPENGICLGLADSFSCPCDGFDPNCS 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 SVEEASDEEPTSGPCTPNPCCHNGTCEISEAYRDTFTGYCKCPRGNGIHCOHNI 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 SVEEASDEEPTSGPCTPNPCCHNGTCEISEAYRDTFTGYCKCPRGNGIHCOHNI 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 NECEVEPCRNKGICITDLVANYSCCEPGEFMRNCQKCSGGLGIEGIIISNOQITASSH 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 NECEVEPCRNKGICITDLVANYSCCEPGEFMRNCQKCSGGLGIEGIIISNOQITASSH 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 RALFGLQKWPYYARLNKKGILNMTAENDRMWIDQINQRMRYGVITOGAKRIGS 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 RALFGLQKWPYYARLNKKGILNMTAENDRMWIDQINQRMRYGVITOGAKRIGS 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 241 PEYIKFYKIAVNDGKTAMVKYKGTNEDMVRGNIDNNTPYANSFTPIKAQYVRLYPQ 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 PEYIKFYKIAVNDGKTAMVKYKGTNEDMVRGNIDNNTPYANSFTPIKAQYVRLYPQ 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 301 VCRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 VCRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 300 ICRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 300 ICRHCTLMELGCELSGCSSEPLGKSGHIDYQITASSIFRTLNMFTWEPKARLD 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 361 KQKYNAMTSGHNDOSOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHM 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 360 KQKYNAMTSGHNDOSOMLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGEHM 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 421 TVYODEKQKDKVFOGNEFDNTHRNKVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 420 TVYODEKQKDKVFOGNEFDNTHRNKVIDPPIYARHRIILPMSWYGRITLASELLGCTEE 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 481 E 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 480 E 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8
US-08-480-229B-21
Sequence 21, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-229B-21

Query Match 64.2% Score 1717; DB 8; Length 321;
Best Local Similarity 98.4%; Pred. No. 2,3e-139;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 158 CSGPLGIEGGIISNOQTASTSTRALFGLOKWPYARLNKGLINAWTAENDRWKMI 217
DB 1 CSGPLGIEGGIISNOQTASTSTRALFGLOKWPYARLNKGLINAWTAENDRWKMI 60
QY 218 QINLQRMKRVTVTOGAKRISPEYIKFYKIAVNDGKTWAMYVKGTNDMVERGND 277
DB 61 QINLQRMKRVTVTOGAKRISPEYIKFYKIAVNDGKTWAMYVKGTNDMVERGND 120
QY 278 NNTPYANSFTPIKAQYVRLYPQVCRRHCTLRMELGCELSGSEPLGKSGHIDYQIT 337
DB 121 NNTPYANSFTPIKAQYVRLYPQVCRRHCTLRMELGCELSGSEPLGKSGHIDYQIT 180
QY 338 ASSIFRTLNMDFTWEPKRLDKQGVNMTSGHNDOSOMLOVDLVPFTKVTGIIITOGA 397
DB 181 ASSIFRTLNMDFTWEPKRLDKQGVNMTSGHNDOSOMLOVDLVPFTKVTGIIITOGA 240
QY 398 KDFGHVQFVGSYKLAISNDEHMTVYQDEKQKRDVFOGNEFDNTHRKNVIDPPIYARHI 457
DB 241 KDXGHVQFVGSYKLAISNDEHMTVYQDEKQKRDVFOGNEFDNTHRKNVIDPPIYARHI 300
QY 458 RILPMSWYGRITLASSELLGCT 478
DB 301 RILPMSWYGRITLASSELLGCT 321

RESULT 9
US-08-659-235B-21
Sequence 21, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quentemus, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-659-235B-21

Query Match 64.2% Score 1717; DB 10; Length 321;
Best Local Similarity 98.4%; Pred. No. 2,3e-139;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 158 CSGPLGIEGGIISNOQTASTSTRALFGLOKWPYARLNKGLINAWTAENDRWKMI 217
DB 1 CSGPLGIEGGIISNOQTASTSTRALFGLOKWPYARLNKGLINAWTAENDRWKMI 60
QY 218 QINLQRMKRVTVTOGAKRISPEYIKFYKIAVNDGKTWAMYVKGTNDMVERGND 277
DB 61 QINLQRMKRVTVTOGAKRISPEYIKFYKIAVNDGKTWAMYVKGTNDMVERGND 120
QY 278 NNTPYANSFTPIKAQYVRLYPQVCRRHCTLRMELGCELSGSEPLGKSGHIDYQIT 337
DB 121 NNTPYANSFTPIKAQYVRLYPQVCRRHCTLRMELGCELSGSEPLGKSGHIDYQIT 180
QY 338 ASSIFRTLNMDFTWEPKRLDKQGVNMTSGHNDOSOMLOVDLVPFTKVTGIIITOGA 397
DB 181 ASSIFRTLNMDFTWEPKRLDKQGVNMTSGHNDOSOMLOVDLVPFTKVTGIIITOGA 240
QY 398 KDFGHVQFVGSYKLAISNDEHMTVYQDEKQKRDVFOGNEFDNTHRKNVIDPPIYARHI 457
DB 241 KDXGHVQFVGSYKLAISNDEHMTVYQDEKQKRDVFOGNEFDNTHRKNVIDPPIYARHI 300
QY 458 RILPMSWYGRITLASSELLGCT 478
DB 301 RILPMSWYGRITLASSELLGCT 321

RESULT 10
US-09-237-981-21
Sequence 21, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quentemus, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29

```

CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2271

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-237-981-21

Query Match          64.2%; Score 1717; DB 16; Length 321;
Best Local Similarity 98.4%; Pred. No. 2.3e-139;
Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 158 CSGPLGEGGIISNOQTASTSTRALFGLOKWPYYARLNKKGLINAMTAENDRRKRWI 217
DB 1 CSGPLGEGGIISNOQTASTSTRALFGLOKWPYYARLNKKGLINAMTAENDRRKRWI 60

QY 218 QINLQRMARVGVITOGAKRIGSPYIKFKYIAVSDGKTWAMKVGNTEDWVFRGND 277
DB 61 QINLQRMARVGVITOGAKRIGSPYIKFKYIAVSDGKTWAMKVGNTEDWVFRGND 120

QY 278 NNPPYANSFPPRIKAQYVRLYPQVCRHRHCTLRMELLGCELSGCSSEPLGKMSGHIQDYQIT 337
DB 121 NNPPYANSFPPRIKAQYVRLYPQVCRHRHCTLRMELLGCELSGCSSEPLGKMSGHIQDYQIT 180

QY 338 ASSIFRTLNMDMTWEPKARLDKQGVNMTSGHNDOSQWLOVDLVPKVTGIIITQGA 397
DB 181 ASSIFRTLNMDMTWEPKARLDKQGVNMTSGHNDOSQWLOVDLVPKVTGIIITQGA 240

QY 398 KDEGHVQFVSGYKLAISNDGEHMTVYODEKQKDKVFOGNFNDTHKKNVIDPPIYARHI 457
DB 241 KDEGHVQFVSGYKLAISNDGEHMTVYODEKQKDKVFOGNFNDTHKKNVIDPPIYARHI 300

QY 458 RILPMSWYGRITTLASELLGCT 478
DB 301 RILPMSWYGRITTLASELLGCT 321

RESULT 11
US-60-207-315-404
; Sequence 404, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

```

```

; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 404
; LENGTH: 449
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(449)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-207-315-404

```

```

Query Match          60.4%; Score 1615; DB 23; Length 449;
Best Local Similarity 70.6%; Pred. No. 2.1e-130;
Matches 305; Conservative 2; Mismatches 7; Indels 118; Gaps 3;

```

```

QY 77 PCTPDPCHNGTCEISEAYRGDTFTIGYVCKCPRGFNGIHCOH----- 118
DB 2 PCTPDPCHNGTCEISEAYRGDTFTIGYVCKCPRGFNGIHCOHRLSHCPALADFLSVPG 61

QY 119 ----- 118
DB 62 KPPLSLIDNAPRPSKNAPIRISTSSGRKVPDLTLREPSDMEHVGPKAAMNTGLAMPVL 121

QY 119 -----NINCEVEPCNKGICIDIVANYSCEGPFGRNCOYKCSGP 161
DB 122 RGGILGQLLRPECGAPINCEVEPCNKGICIDIVANYSCEGPFGRNCOYKCSGP 181

QY 162 LGIEGGIISNOQTASTSTRALFGLOKWPYYARLNKKGLINAMTAENDRRKRWI----- 214
DB 162 LGIEGGIISNOQTASTSTRALFGLOKWPYYARLNKKGLINAMTAENDRRKRWI----- 241

QY 215 ---RWIQINLQRMARVGVITOGAKRIGSPYIKFKYIAVSDGKTWAMKVGNTEDM 270
DB 242 CHLGWPDQINLQRMARVGVITOGAKRIGSPYIKFKYIAVSDGKTWAMKVGNTEDM 301

QY 271 VFRGNDNNTPYANSFPPRIKAQYVRLYPQVCRHRHCTLRMELLGCELSGCSSEPLGKMSGH 330
DB 302 VFRGNDNNTPYANSFPPRIKAQYVRLYPQVCRHRHCTLRMELLGCELSGCSSEPLGKMSGH 361

QY 331 IDQYQITASSIFRTLNMDMTWEPKARLDKQGVNMTSGHNDOSQWLO----- 380
DB 362 IDQYQITASSIFRTLNMDMTWEPKARLDKQGVNMTSGHNDOSQWLOFTLLISVEYSG 421

QY 381 ---VDLVPKVT 390
DB 422 TALALMVPTAAT 433

RESULT 12
US-60-207-315-523
; Sequence 523, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 523
; LENGTH: 362
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-315-523

```

Query Match	52.6%;	Score 1406.5;	DB 23;	Length 362;
Best Local Similarity	88.8%;	Pred. No. 1.4e-112;		
Matches 262; Conservative	3;	Mismatches 7;	Indels 23;	Gaps 2;

QY	119	NINNEVEPKNGICITD VANSCECEPEFPMGRNCOYKCSGPIGIEGIIISNOQITASS	178
	:		
Db	52	DINECEVEPKNGICITD VANSCECEPEFPMGRNCOYKCSGPIGIEGIIISNOQITASS	111
QY	179	THRALFGLQWYDYV ARLNKGLINAMTAENDRMK-----RWIOINQRMRY	227
	:		
Db	112	THRALFGLQWYDYV ARLNKGLINAMTAENDRMWPWIOELRGCHLGPWQD INQRMRY	171
QY	228	TGVTGQARKIGSEPEYIKFKI YNSDKGTWAMYVVKGNEDMVRGINDNPTYANSFT	287
Db	172	TGVTGQAKKIGSEPEYIKFKI YNSDKGTWAMYVVKGNEDMVRGINDNPTYANSFT	231
QY	288	PKIAQYVRLPYOVCRRHCTLRNELLGCELSGCSEPLMGSGHIDVOYITASSIFRTLNM	347
Db	232	PKIAQYVRLPYOVCRRHCTLRNELLGCELSGCSEPLMGSGHIDVOYITASSIFRTLNM	291
QY	348	DMFTWBRKARLIDKOGVNAWMTSGHNDQSOHWI-----VDDLVP TKYT	390
Db	292	DMFTWBRKARLIDKOGVNAWMTSGHNDQSOHWI PTLISYEVSCTLALMVTAAFT	346

Query Match	52.58;	Score 1405.5;	DB 23;	Length 311;
Best Local Similarity	89.18;	Pred. No. 1.4e-112;		
Matches 262;	Conservative	2;	Mismatches 7;	Indels 23;
				Gaps 2

Qy	120	INEVEBPCKNNGICITDVLVANTSCCEPGEPMGNCYKSCGLGEGGIIISMOQTAST	179
Pb	2	INEVEBPCKNNGICITDVLVANTSCCEPGEPMGNCYKSCGLGEGGIIISMOQTAST	61
Qy	180	HRALEFGLOKMYPYARLNKKGGLNMTAENBWK-----RWOTINLOKMKMT	228
Pb	62	HRALEFGLOKMYPYARLNKKGGLNMTAENBWKPIYOELRCHGLGPMDOINLOKMKMT	121
Qy	229	GVTGQAKRIGSPETIKFYKILATSNQKTMAMKXNGTIEDWYFNGKNDNNTPYANSTP	288
Pb	122	GVTGQAKRIGSPETILSYKILATSNQKTMAMKXNGTIEDWYFNGKNDNNTPYANSTP	181
Qy	289	PIKAQYVRLYPQVCRHRCTLMELLCGELSGSEPLGKMSGHIDYQITASSIFPTLMD	348
Pb	182	PIKAQYVRLYPQVCRHRCTLMELLCGELSGSEPLGKMSGHIDYQITASSIFPTLMD	241
Qy	349	MTTEPRKARLDKQKVNAVNTSGHNDOSOMLO-----VDLVTPTKVT	390
Pb	242	MTTEPRKARLDKQKVNAVNTSGHNDOSOMLOPTLLISVEGSTATLAWYPTAT	295

```

RESULT 14
US-09-582-340-4
: Sequence 4, Application US/09582340
: GENERAL INFORMATION:
: APPLICANT: INSERM
: APPLICANT: INSTITUT CURIE
: APPLICANT: CNRS
: TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants
: TITLE OF INVENTION: thereof
: FILE REFERENCE: Lactadherin
: CURRENT APPLICATION NUMBER: US/09/582,340
: CURRENT FILING DATE: 2000-09-15
: NUMBER OF SEQ. ID NOS.: 6
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 463
: TYPE: PRT
: ORGANISM: mammalian
US-09-582-340-4

```

Query Match	43.2%	Score 1156;	DB 19;	Length 463;
Best Local Similarity	48.5%	Pred. No. 6.7e-91;		
Matches 222; Conservative	72;	Mismatches 142;	Indels 22;	Gaps 8;

[illegible]

RESULT 15
US-09-237-981-29
Sequence 29, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zapancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:52 ; Search time 28 Seconds
(without alignments)
123.138 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513
Perfect score: 2675
Sequence: 1 MKRSVAVMLVGLSLGVPQF.....MSWYGRITLASLGLCTHEE 481

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41833 seqs, 7168127 residues

Total number of hits satisfying chosen parameters: 41833

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	466.5	17.4	923	5	US-09-583-638-2
2	462.5	17.3	957	1	PCT-US00-35017A-1285
3	453	16.9	931	5	US-09-583-638-4
4	287	10.7	723	5	US-09-828-366-21
5	270.5	10.1	188	1	PCT-US01-01302-109
6	265	9.9	226	1	PCT-US01-01302-92
7	261	9.8	1964	5	US-09-467-997-1
8	243.5	9.1	383	1	PCT-US01-11988-770
9	243.5	9.1	383	5	US-09-833-245-770
10	242.5	9.1	310	1	PCT-US01-01302-103
11	234	8.7	734	1	PCT-US01-09226-65
12	191	7.1	4044	6	US-60-248-505-791
13	188	7.0	904	1	PCT-US00-35017A-1079
14	184.5	6.9	3014	1	PCT-US01-10484-63
15	184	6.9	213	1	PCT-US01-01302-90
16	184	6.9	213	1	PCT-US01-01302-90
17	182	6.8	379	5	US-09-808-689-4
18	182	6.8	379	5	US-09-423-844-4
19	182	6.8	379	5	US-09-403-296A-4
20	182	6.8	379	5	US-09-380-139A-4
21	178	6.7	501	1	PCT-US01-04098A-3346
22	178	6.7	4544	1	PCT-US01-04098A-1753
23	176	6.6	4545	5	US-09-750-972-2
24	171	6.4	1139	1	PCT-US01-01302-99
25	171	6.4	1139	1	PCT-US01-01332-873
26	170.5	6.4	534	5	US-09-804-156-14
27	170	6.4	576	1	PCT-US01-04098A-3820

28	170	6.4	576	1	PCT-US01-04098A-3821	Sequence 3821, Ap
29	170	6.4	603	1	PCT-US01-04098A-1853	Sequence 1853, Ap
30	170	6.4	650	1	PCT-US01-04098A-1852	Sequence 1852, Ap
31	168	6.3	122	1	PCT-US01-01302-97	Sequence 97, Appl
32	166	6.2	201	1	PCT-US01-10484-74	Sequence 74, Appl
33	159.5	6.0	509	5	US-09-423-844-315	Sequence 315, App
34	159.5	6.0	509	5	US-09-380-139A-314	Sequence 314, App
35	158.5	5.9	835	6	US-60-283-145-43	Sequence 43, Appl
36	149	5.6	415	5	US-09-802-704-7	Sequence 627, App
37	149	5.6	439	1	PCT-US01-11988-627	Sequence 627, App
38	149	5.6	439	1	PCT-US01-11988-628	Sequence 628, App
39	149	5.6	439	5	US-09-802-704-4	Sequence 4, Appl1
40	149	5.6	439	5	US-09-833-245-627	Sequence 627, App
41	149	5.6	439	5	US-09-833-245-628	Sequence 628, App
42	146.5	5.5	156	1	PCT-US01-01302-95	Sequence 95, Appl
43	146	5.5	622	1	PCT-US01-04098A-3708	Sequence 3708, Ap
44	145	5.4	776	1	PCT-US01-04098A-1740	Sequence 1740, Ap
45	145	5.4	961	5	US-09-657-472-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-583-638-2
Sequence 2, Application US/09583638
GENERAL INFORMATION:
APPLICANT: KLAGSBRUN, MICHAEL
APPLICANT: SOKER, SHAY
APPLICANT: MIAO, HUA-QUN
APPLICANT: TARASHIMA, SEIJI
TITLE OF INVENTION: NEUROPLAINS AND USE THEREOF IN METHODS FOR
FILE REFERENCE: 701039-48800
CURRENT APPLICATION NUMBER: US/09/583,638
CURRENT FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: PCT/US98/26127
PRIOR FILING DATE: 1998-12-09
PRIOR APPLICATION NUMBER: 60/069,155
PRIOR FILING DATE: 1997-12-09
PRIOR APPLICATION NUMBER: 60/069,687
PRIOR FILING DATE: 1997-12-12
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 923
TYPE: PRT
ORGANISM: Homo sapiens
US-09-583-638-2

Query Match 17.4% Score 466.5; DB 5; Length 923;
Best Local Similarity 32.1% Pred No. 1.3e-34;
Matches 159; Conservative 59; Mismatches 210; Indels 67; Gaps 23;

QY	12	GLISLGVPGFGKGDIDPN-PCENGICILPLAVG-SFSCPCPDGFTDPCSSVY-EVASD	68
DB	133	GRSIVTEIFKRGPECSQVYTPSGVTKSPPEKYPNSLETCTIVFAPKMSEILFESEF	192
QY	69	EEEPISAPCTPNNPCNNGCTEISAVRGDTF-----IGYVC--KCP--RCNGNHQ	117
DB	193	DLEPDSNPP-----GGMFCRYDLREITWDEFPDVGPHRGYCOQKTPGRIRSSGILSM	245
QY	118	HNINCEVEPCNGKICIDLVANYS-C-EGGEFMGRNCOYKSGSPIGEGILSNOIT	175
DB	246	VFTYDSAI--AKEG-----FSANISVLOSYSF-----DFKCEALGMSGETHSQIT	292
QY	176	ASSTRALFGLQKWPPYARLNKGLINAWTAENDRKRWIQINLORKKRVTVGITOGA	235
DB	293	ASSQVST-----NMSAEKSRINLYPE--NGWTPGE--DSYREMIQYDGLLRVTAVGTOGA	344
QY	236	--KRIGSEPIYKFKIANSNDGKTWAMKYVGTEDMVFRRGINDNNTPYANSFPPIKAQ	293

```
Db 345 ISKETKKKKYKTYKIDVSSNGEDW--ITIEGKNKPVLFQGNTPEDVYVAVPEPKLITR 402
QY 294 YVRLYPQVCRHCTLRMELLGCELSG--CSEPLGKMSGHIQDQIATASIFRLNMDMT 351
Db 403 FVRKIPATWETGISMREYVGTCTIDYPCSGMLGWSGLISDQITSS-----NOGDRN 456
QY 352 WEPKARL--DKOGKVAMT---SGHNDOSQMLQVDLLVTKYTGITTOGAKDFGHVOFY 406
Db 457 WMPENILVYSRSG---WALPAPRHSYINEMWIDIDGEEKIYRGITLIOGKIREKNKVF 512
QY 407 GSYLAVSNDGEHWTYQDEKOKKQVFCQGFNDQTHRKRVNDPPIYARHRIPLP--WSWY 465
Db 513 RKFKIGYSNNGSDMKMIMDSKRAKSFEGNNNDPELRTF--PALSTRPFRIRYPERATH 571
QY 466 GRITLASELCTEE 480
Db 572 GGLGLRMELLGCEVE 586

RESULT 2
PCT-US00-35017A-1285
; Sequence 1285, Application PC/TUS0035017A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: PCT/US00/35017A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1285
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-35017A-1285

Query Match 17.3%; Score 462.5; DB 1; Length 957;
Best Local Similarity 31.9%; Pred. No. 3.2e-34;
Matches 158; Conservative 60; Mismatches 210; Indels 67; Gaps 23;

QY 12 GLSLGVQFGKGDICDPN-PCENGICLPGLAVG-SFSCPCPDGFTDPNCSYV-EVYASD 68
Db 167 GFSTRYELFKRGPCSQNTTPSGVIKSPGFPEKYPMSLCTIYVAPKSEILIDPESEF 226
QY 69 EEPETSAQPCPTPNCHNGGTCETSEAYRGDTF-----IGYVC--KCP---RGFNGIHQ 117
Db 227 DLEPDSNPP-----GGMFCRYDRLEIMGFPDVGPHIGRYCGOKTPGRIRSSGILSM 279
QY 118 HNIECEVEPCXNGICITDLVANYSC--ECPGEFMGRNOCYKCSGPIGTGIGIISNOQT 175
Db 280 VFYIDSAI--AKESG-----FSANISVLQSSVSE-----DFKCEALGMSGSIHSDQIT 326
QY 176 ASSTHRAFLFGLOKWPYARLNNKGLINAMTAAENDRMKRWIOLNQRKRVYGVITOGA 235
Db 327 ASSQYST-----MMSAERSRLNPE--NGWTPE--DSYREMIQVDGLLRFVYAVGTQGA 378
QY 236 --KRIGSEPIKFKYKLAIVSNDGKTWAMYKVGKNEDMVFEGNIDNNTPYANSTPPIKAQ 293
Db 379 ISKETKKKKYKTYKIDVSSNGEDW--ITIEGKNKPVLFQGNTPEDVYVAVPEPKLITR 436
QY 294 YVRLYPQVCRHCTLRMELLGCELSG--CSEPLGKMSGHIQDQIATASIFRLNMDMT 351
Db 437 FVRKIPATWETGISMREYVGTCTIDYPCSGMLGWSGLISDQITSS-----NOGDRN 490
QY 352 WEPKARL--DKOGKVAMT---SGHNDOSQMLQVDLLVTKYTGITTOGAKDFGHVOFY 406
Db 491 WMPENILVYSRSG---WALPAPRHSYINEMWIDIDGEEKIYRGITLIOGKIREKNKVF 546
QY 407 GSYLAVSNDGEHWTYQDEKOKKQVFCQGFNDQTHRKRVNDPPIYARHRIPLP--WSWY 465
```

```
Db 547 RKFKIGYSNNGSDMKMIMDSKRAKSFEGNNNDPELRTF--PALSTRPFRIRYPERATH 605
QY 466 GRITLASELCTEE 480
Db 606 GGLGLRMELLGCEVE 620

RESULT 3
US-09-583-638-4
; Sequence 4, Application US/09583638
; GENERAL INFORMATION:
; APPLICANT: KIAGSBRUN, MICHAEL
; APPLICANT: SOKER, SHAY
; APPLICANT: MIAO, HUA-QUAN
; APPLICANT: TAKASHIMA, SEIJI
; TITLE OF INVENTION: NEUROFILINS AND USE THEREOF IN METHODS FOR
; TITLE OF INVENTION: DIAGNOSIS AND PROGNOSIS OF CANCER
; FILE REFERENCE: 701039-48800
; CURRENT APPLICATION NUMBER: US/09/583,638
; CURRENT FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US98/26127
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 60/069,155
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 60/069,687
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-583-638-4

Query Match 16.9%; Score 453; DB 5; Length 931;
Best Local Similarity 35.4%; Pred. No. 2.2e-33;
Matches 120; Conservative 59; Mismatches 122; Indels 38; Gaps 15;

QY 156 YKCSGGLGIEGIIISNOQTASTHRALEFGLOKWPYARLNNKGLINAMTAAENDRMKR 215
Db 275 FQCNVPLGESSIRINEQSASST---YSDGRWTPQOSRLH--GDDGWT--PILDSNKE 327
QY 216 WQIINLQRMARYGVITOGA--KRIGSEPIKFKYKLAIVSNDGKTWAMYKVGKNEDMVF 273
Db 328 YLQVDLRFMTLALATQGAISRETQNGYVYKYLEVSTNEDMVMYR-HGKNH-KVFQ 385
QY 274 GNIDNTPYANSTPPIKAQVYRLYPQVCRHCTLRMELLGCELSG--CSEPLGKMSGHI 331
Db 386 ANNDATEVVLNKLHAPRLTRFVIRIPQTHNSGIALRLLEFGCRVYDAPCSNMLGMLSLGI 445
QY 332 QYQYTASSIFFTLMDMTWEPKARL--DKOG---KVNMTSGHNDOSQMLQVDLLVP 386
Db 446 ADSQISASS-----TQELVMSPSAARLVSSRSGWFPRIPOAPR---EEMLOVDLCTP 495
QY 387 TQVTGIIITOGAKDFGHV-----OFVGSYKLAIVSNDGEHWTYQDEKOKKQVFCQGFND 440
Db 496 KTVKCVITIOGAGSGSITAVEARAFVRKFKVYSYLSNGKDWELIDQPRIOQKLFEGNNHY 555
QY 441 DTHRKNVNDPPIYARHRIPLP--WSWYGRITLASELCTEE 477
Db 556 DTPDIRRPD-PIPAQYVRYVPERMSPAG-IGMRLEVLGC 592
```

```
RESULT 4
US-09-828-366-21
; Sequence 21, Application US/09828366
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
```



```

1  APPLICANT: Klein, Robert D.
2  APPLICANT: Napier, Mary
3  APPLICANT: Wood, William I.
4  APPLICANT: Yuan, Jean
5  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
6  TITLE OF INVENTION: CELL GROWTH
7  FILE REFERENCE: P1694R1C1
8  CURRENT APPLICATION NUMBER: US/09/828,366
9  CURRENT FILING DATE: 2001-04-05
10 Prior filing data removed - refer to PALM or file wrapper
11 NUMBER OF SEQ ID NOS: 29
12 SEQ ID NO 21
13 LENGTH: 723
14 TYPE: prt
15 ORGANISM: Homo Sapien
16 US-09-828-366-21

```

Query Match	10.7%	Score 287;	DB 5;	Length 723;
Best Local Similarity	35.0%	Pred. No. 1.7e-18;		
Matches 57;	Conservative 12;	Mismatches 48;	Indels 46;	Gaps 4

QY	24	DIDPNPNCENGJCTCPGLAVSFSCECDGDTTPNCSSVPEVAASDEEPTSTAGSGPTMPC	83
		:	
Db	330	DECDPSPCKNGGSCITD--LENSYSCTCPPGFYKICE--	375
		-----LSAMTCADPDC	
QY	84	HNGSTEISE-----AYRG-----	111
Db	376	FNGRCSDSPDGGISCRCPYVGSQNCNEKRTDYCSSPSCNGAKCYDGLDADATLCRCQAGF	433
QY	112	NGTHCHQHNINECEVPCCKNGGCTCTDLYANVNSCECPGPFPMGRNC	154
		:	
Db	436	SGRHCDNDVDDCASSPCAGGCTCRGVVDSFCTCPGPTGRNC	478

```

RESULT      5
PCT-US01-01302-109
? Sequence 109, Application PC/TUS0101302
? GENERAL INFORMATION:
? APPLICANT: Human Genome Sciences, Inc., et al.
? TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
? FILE REFERENCE: PM707PCT
? CURRENT APPLICATION NUMBER: PCT/US01/01302
? NUMBER OF SEQ ID NOS: 192
? SOFTWARE: Patentl Ver. 2.0
? SEQ ID NO 109
? LENGTH: 188
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (97)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (187)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (188)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-109

```

```

Query Match      10.1%  Score 270.5; DB 1; Length 188;
Best Local Similarity 35.6%  Pred. No. 9,5e-18;
Matches 56; Conservative 18; Mismatches 54; Indels 33; Gaps 7

QY      23 GDICDP-----NPGENGIGLPLGIAVSFSCRCPPGFTTPNCCSVVEAVASDEEPTS 74
      ||  ||  ||: ||  ||: ||  ||: ||  ||: ||  ||  ||
DB      28 GLFDQDQLYNCTHSPCKNGATC-SNSGGRSTYCTCRGTYGYDC---ELTSE----- 77
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

QY      75 AGPTPTPCNCHNGGICETSEAYRGDTFTGYVCCPPRGFNGIHCOHININECEVPPCKNGIC 134
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

```

Db 78 ---CDSNPCCANGSCCKDDED-----GYXLCIPGGYGGHLCHESTPLSCADSPCCNNGSC 127
 Qy 135 TD--LVANSCCEGPGEFMRNCOYK--CSGPLDIEGIIISL 172
 Db 128 RERMGANNTACGCPFFMTFSNCEKIVDRCTISPCANGGCLNR 170

```

RESULT 6
PCT-US01-01302-92
: Sequence 92, Application PC/TUS0101302
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc., et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT20/PCT
: CURRENT APPLICATION NUMBER: PCT/US01/01302
: NUMBER OF SEQ ID NOS: 192
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 92
: LENGTH: 226
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (133)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-92

```

		9.9%	Score 265;	DB 1:	Length 226;
	Query Match				
	Best Local Similarity	32.7%	Pred No. 3.8e-17;		
	Matches	54;	Conservative	19;	Mismatches 58; Indels 34; Gaps 5;
Oy	26	CDNPCCENGICIPGLAVGSFSCEDGDTDFDNCSSVVEVADEDEPTTSAGPCTPNPCAN	85		
		: : :			
Dd	28	CQSNPLCNNAVCD--QYSGFMKCKCPGSLGRCKNND-----ECLISQPCKN	73		
Oy	86	GCTCEISEAYRGDTFTIGLYVCKCPRGFNHICHQHINICEVEPPCKNGICTIDLVANYSCEC	145		
		: :			
Dd	74	GATKC-----DOANSFRICACAGTGTSHCELINIECOSMPCNRQATCYVDLNSYSCRC	126		
Oy	146	PGEFMGRNQ-----YKCSPLG--TEGGILINQQTAS	177		
		: : : :			
Dd	127	QPGEFGAKRCEETEGSTGFNDLFVESLGTYGVMDGMPSLHALTCCT	171		

```

Oy      26  CDPMPCKNGGICLPLGLAVGSFSCGCPDGGFTDPNCSVVEVASDEDEPTSA G PCTPMPCHN 85
      | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      28  COSMPCLNNAVCE--QVGSGMCKCPGFLGTRCKNVD-----ECLSQPCAN 73
      | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Oy      86  GGTCEISEAVRGDTFTGYCKCPRGFGNHCOHINCECEVPCKNGGICHTDLYANTSCBC 14
      | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db      74  GATCK-----DGANSFRLCAAGTGSCHCLININECOSMPCNMOATCVDLEINRYCKC 12
      | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Oy      146  PGEFMGRNCO-----YKOSGRLG--EEGIIINQQTAS 177
      | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db      127  QPGEISGXRCETEGSTGFNDEFEVSYGIYGVMLDGMPLPSLHALTCT 171
      | | | | | : | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-467-997-1
: Sequence 1, Application US/09467997
: GENERAL INFORMATION:
: APPLICANT: Kitajewski, Jan
: APPLICANT: Uytendaele, Hendrik
: TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
: FILE REFERENCE: 53863-A-PCT-US
: CURRENT APPLICATION NUMBER: US/09/467,997
: CURRENT FILING DATE: 1999-12-20
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 1964
: TYPE: PRT
: ORGANISM: mouse
US-09-467-997-1

```

```

      RESULT          7
US-09-467-997-1
; Sequence 1, Application US/09467997
; GENERAL INFORMATION:
; APPLICANT: KitaJewski, Jan
; APPLICANT: Uytendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCR-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: PRT
; ORGANISM: mouse
US-09-467-997-1

Query Match             9.8%; Score 261; DB 5; Length 1964;
Best Local Similarity   29.9%; Pred. No. 1.4e-15;
Matches 60; Conservative 19; Mismatches 60; Indels 62; Gaps

QY      24 DICPNPCNGCGICL-----PGL-----AVGS 45
       |  ||| |||||
      155 DFCGANGCANGCVLATVYDIOCRCPGGHGTTCERDINCFLEPGCPQGTSCHTNIGS 214
```

```

QY      24 DICDPNCCNGICL-----PGL-----AYGS 45
          | | | | | | | | | |
Db      155 DFCSANPCANGVCLATYPIQICRCPPEGSHTCERDINECFLEPFCPOGMSCHNTLGS 214

```


OY 122 ECEVEPCCKNGICTD-LVANYSCC-PGEFMRNQ 155
 Db 1406 RCANGVCCKNGICVNLIGGFHCVCPPGEYERPC 1441

RESULT 15
 PCT-US01-01302-90
 ; Sequence 90, Application PC/TUS0101302
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc., et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT207PCT
 ; CURRENT APPLICATION NUMBER: PCT/US01/01302
 ; CURRENT FILING DATE: 2001-01-17
 ; NUMBER OF SEQ ID NOS: 192
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 90
 ; LENGTH: 213
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (17)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (179)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (195)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (205)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; PCT-US01-01302-90

Query Match 6.9%; Score 184; DB 1; Length 213;
 Best local Similarity 26.5%; Pred. No. 7.6e-10;
 Matches 44; Conservative 22; Mismatches 64; Indels 36; Gaps 6;
 OY 3 RSVAVMLVGLSGVDFGKIDIC-----PNCENGICLPGLAVGSFS---CECP 51
 Db 29 KNINYLIFLSFLIYIKSFCKNNTRCLISNCQNNSTC-----KDFSKDNDSCS 82
 OY 52 D-GTDPNCSVVEVASDEEPTAGCTPNPCHNGSTCEISEAYRGDTFYGVCCKPR 109
 Db 83 DTANNDDKDDNMKD-----PCFSNRCGGSATCVNTPGER-----SFLCKCP 125
 OY 110 GFNGIHQOHINCEVEPCCKNGICTDLVANYSCC-PGEFMRNQ 155
 Db 126 GYSGTICETITGSCGKNSQHGICHODPIYVVICIPAGYAGRCE 171

Search completed: May 23, 2001, 06:13:54
 Job time: 148 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:18 ; Search time 38.64 Seconds
(without alignments)
239.141 Million cell updates/sec

Title: US-09-237-981-14_COPY_33_513

Perfect score: 2675
Sequence: 1 MKRSVAVMLLVGLSLGVPOF.....MSWYGRITLSELCTEEH 481

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2-6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2-6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2-6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2-6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2-6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2-6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2675	100.0	513	2	US-08-480-229C-14
2	2675	100.0	513	2	US-08-659-235C-14
3	2519.5	94.2	480	2	US-08-480-229C-10
4	2519.5	94.2	480	2	US-08-659-235C-10
5	1717	64.2	321	2	US-08-480-229C-21
6	1717	64.2	321	2	US-08-659-235C-21
7	1144	42.8	463	2	US-08-162-402B-9
8	1133.5	42.4	221	2	US-08-480-229C-29
9	1133.5	42.4	221	2	US-08-659-235C-29
10	1120.5	41.9	465	2	US-08-162-402B-6
11	1020.5	38.1	387	2	US-08-162-402B-6
12	948	35.4	320	2	US-08-162-402B-20
13	948	35.4	320	2	US-08-659-235C-20
14	682.5	25.5	2183	3	US-08-746-111-5
15	670	25.0	2319	1	US-08-212-133A-8
16	670	25.0	2319	1	US-08-474-503-6
17	670	25.0	2319	1	US-08-670-707A-6
18	670	25.0	2319	4	US-09-037-601-6
19	670	25.0	2319	5	PCT-US94-13200-6
20	653	24.4	2332	1	US-08-276-594A-2
21	650	24.3	2351	6	5422260-1
22	649	24.3	1471	1	US-08-683-839B-3
23	649	24.3	1661	2	US-08-882-083-2
24	649	24.3	1661	2	US-08-558-107-2
25	649	24.3	1661	4	US-09-243-539-2
26	649	24.3	2332	1	US-07-864-004B-4
27	649	24.3	2332	1	US-08-251-937A-4

28	649	24.3	2332	1	US-08-212-133A-2	Sequence 2, Appl1
29	649	24.3	2332	1	US-08-474-503-2	Sequence 2, Appl1
30	649	24.3	2332	2	US-08-670-707A-2	Sequence 2, Appl1
31	649	24.3	2332	4	US-09-037-601-2	Sequence 2, Appl1
32	649	24.3	2332	5	PCT-US93-03275-4	Sequence 4, Appl1
33	649	24.3	2332	5	PCT-US94-13200-2	Sequence 2, Appl1
34	649	24.3	2351	1	US-08-121-202-2	Sequence 2, Appl1
35	649	24.3	2351	1	US-08-366-851A-2	Sequence 2, Appl1
36	643	24.0	217	2	US-07-607-538C-3	Sequence 3, Appl1
37	643	24.0	217	2	US-08-162-402B-3	Sequence 3, Appl1
38	643	24.0	218	1	US-07-607-538C-2	Sequence 2, Appl1
39	643	24.0	218	2	US-08-162-402B-2	Sequence 2, Appl1
40	643	24.0	2351	6	5171844-2	Sequence 39, Appl1
41	639	23.9	1443	2	US-08-670-707A-39	Sequence 39, Appl1
42	639	23.9	1443	4	US-09-037-601-39	Sequence 37, Appl1
43	639	23.9	2133	2	US-08-670-707A-37	Sequence 37, Appl1
44	639	23.9	2133	4	US-09-037-601-37	Sequence 13, Appl1
45	501	18.7	157	2	US-08-162-402B-13	

ALIGNMENTS

RESULT 1
US-08-480-229C-14
; Sequence 14, Application US/08480229C
; Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quaternous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennle & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229C-14

Query Match 100.0%; Score 2675; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 2.8e-210; Indels 0; Gaps 0;
Matches 481; Conservative 0; Mismatches 0;

```

QY 1 MKRVAWMLVGLSLGVPQFGKGDICDPNCEGNGICLPLGLAVGSFCECPDGTDPNCS 60
   |||
   |||
   |||
Db 33 MKRVAWMLVGLSLGVPQFGKGDICDPNCEGNGICLPLGLAVGSFCECPDGTDPNCS 92
   |||
   |||
   |||
QY 61 SVVEVASDEEPTSAQPCPNPCNHGTCISEAYRGDTFTGYVCKCPRGFNGIHCOHNI 120
   |||
   |||
   |||
Db 93 SVVEVASDEEPTSAQPCPNPCNHGTCISEAYRGDTFTGYVCKCPRGFNGIHCOHNI 152
   |||
   |||
   |||
QY 121 NECEVEPCCKNGGICITDLVANYSCCEPGEPMGRNCOYKCSGPLGIGGIIISNOQITASSTH 180
   |||
   |||
   |||
Db 153 NECEVEPCCKNGGICITDLVANYSCCEPGEPMGRNCOYKCSGPLGIGGIIISNOQITASSTH 212
   |||
   |||
   |||
QY 181 RALFGLQKWPPYARLNLKKGILINAMTAENDRMRWIOIOLQRMRYTVITOGAKRIGS 240
   |||
   |||
   |||
Db 213 RALFGLQKWPPYARLNLKKGILINAMTAENDRMRWIOIOLQRMRYTVITOGAKRIGS 272
   |||
   |||
   |||
QY 241 PEYIKFYKIAVNDGKTAMKYKGTNEDMVRGNIIDNNTPRANSFTPPIKAQYVRLYPQ 300
   |||
   |||
   |||
Db 273 PEYIKFYKIAVNDGKTAMKYKGTNEDMVRGNIIDNNTPRANSFTPPIKAQYVRLYPQ 332
   |||
   |||
   |||
QY 301 VCRHCTLRMELLGCELSGCSSEPLGKMSGHIDYQITASSIFRTLNMDMTWEPKARLD 360
   |||
   |||
   |||
Db 333 VCRHCTLRMELLGCELSGCSSEPLGKMSGHIDYQITASSIFRTLNMDMTWEPKARLD 392
   |||
   |||
   |||
QY 361 KQGVNAMTSGHNDOSQWLQVDLLVPTKVTGIITOGAKDFGHVQFVGSYKLAYSNDGEHM 420
   |||
   |||
   |||
Db 393 KQGVNAMTSGHNDOSQWLQVDLLVPTKVTGIITOGAKDFGHVQFVGSYKLAYSNDGEHM 452
   |||
   |||
   |||
QY 421 TVYODEKORKDKVFOGNFNDTHRKNNVIDPPIYARHIRILPMSWGRTILASLCTEE 480
   |||
   |||
   |||
Db 453 TVYODEKORKDKVFOGNFNDTHRKNNVIDPPIYARHIRILPMSWGRTILASLCTEE 512
   |||
   |||
   |||
QY 481 E 481
   |||
   |||
   |||
Db 513 E 513
   |||
   |||
   |||

```

```

RESULT 2
US-08-659-235C-14
; Sequence 14, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659, 235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELE: 66141 Pennie

```

```

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235C-14

```

```

Query Match 100.0%; Score 2675; DB 2; Length 513;
Best Local Similarly 100.0%; Pred. No. 2.8e-210;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MKRVAWMLVGLSLGVPQFGKGDICDPNCEGNGICLPLGLAVGSFCECPDGTDPNCS 60
   |||
   |||
   |||
Db 33 MKRVAWMLVGLSLGVPQFGKGDICDPNCEGNGICLPLGLAVGSFCECPDGTDPNCS 92
   |||
   |||
   |||
QY 61 SVVEVASDEEPTSAQPCPNPCNHGTCISEAYRGDTFTGYVCKCPRGFNGIHCOHNI 120
   |||
   |||
   |||
Db 93 SVVEVASDEEPTSAQPCPNPCNHGTCISEAYRGDTFTGYVCKCPRGFNGIHCOHNI 152
   |||
   |||
   |||
QY 121 NECEVEPCCKNGGICITDLVANYSCCEPGEPMGRNCOYKCSGPLGIGGIIISNOQITASSTH 180
   |||
   |||
   |||
Db 153 NECEVEPCCKNGGICITDLVANYSCCEPGEPMGRNCOYKCSGPLGIGGIIISNOQITASSTH 212
   |||
   |||
   |||
QY 181 RALFGLQKWPPYARLNLKKGILINAMTAENDRMRWIOIOLQRMRYTVITOGAKRIGS 240
   |||
   |||
   |||
Db 213 RALFGLQKWPPYARLNLKKGILINAMTAENDRMRWIOIOLQRMRYTVITOGAKRIGS 272
   |||
   |||
   |||
QY 241 PEYIKFYKIAVNDGKTAMKYKGTNEDMVRGNIIDNNTPRANSFTPPIKAQYVRLYPQ 300
   |||
   |||
   |||
Db 273 PEYIKFYKIAVNDGKTAMKYKGTNEDMVRGNIIDNNTPRANSFTPPIKAQYVRLYPQ 332
   |||
   |||
   |||
QY 301 VCRHCTLRMELLGCELSGCSSEPLGKMSGHIDYQITASSIFRTLNMDMTWEPKARLD 360
   |||
   |||
   |||
Db 333 VCRHCTLRMELLGCELSGCSSEPLGKMSGHIDYQITASSIFRTLNMDMTWEPKARLD 392
   |||
   |||
   |||
QY 361 KQGVNAMTSGHNDOSQWLQVDLLVPTKVTGIITOGAKDFGHVQFVGSYKLAYSNDGEHM 420
   |||
   |||
   |||
Db 393 KQGVNAMTSGHNDOSQWLQVDLLVPTKVTGIITOGAKDFGHVQFVGSYKLAYSNDGEHM 452
   |||
   |||
   |||
QY 421 TVYODEKORKDKVFOGNFNDTHRKNNVIDPPIYARHIRILPMSWGRTILASLCTEE 480
   |||
   |||
   |||
Db 453 TVYODEKORKDKVFOGNFNDTHRKNNVIDPPIYARHIRILPMSWGRTILASLCTEE 512
   |||
   |||
   |||
QY 481 E 481
   |||
   |||
   |||
Db 513 E 513
   |||
   |||
   |||

```

```

RESULT 3
US-08-480-229C-10
; Sequence 10, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quettermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480, 229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229C-10

```

```

Query Match          94.2%; Score 2519.5; DB 2; Length 480;
Best Local Similarity 94.2%; Pred. No. 1.2e-197;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

```

```

QY 1 MKRVAWLLVGLSLGVPOFGKGDICDPNCEGICLPLGLAVSFCSPGDTFDPNCS 60
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 1 MKHLVAMLLVGLSLGVPOFGKGDICDPNCEGICLPLGLAVSFCSPGDTFDPNCS 60
QY 61 SVVEVASDEEPTSGAGPCIPNCHNGGTCISEAYRGDTFGLGYCKCPRGNGIHCOHNI 120
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 61 SVVEVASDEEPTSGAGPCIPNCHNGGTCISEAYRGDTFGLGYCKCPRGNGIHCOHNI 120
QY 121 NECEVPECKNGGICITDLVANYSCGCEPGEFMRNCOYKCSGPLGIEGIIISNOQITASSSTH 180
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 121 NECEVPECKNGGICITDLVANYSCGCEPGEFMRNCOYKCSGPLGIEGIIISNOQITASSSTH 180
QY 181 RALFGLQKWYPYARLNKKGLINAMTAENDRMKRWIOLNORRMRYGVITOGAKRIGS 240
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 181 RALFGLQKWYPYARLNKKGLINAMTAENDRMW-IQINLORRMRYGVITOGAKRIGS 239
QY 241 PEYIKFYKIAVSNCKTAMKVKGTNEDMVRGINDNNTFYANSFTPIKAQVRLYPQ 300
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 240 PEYIKFYKIAVSNCKTAMKVKGTNEDMVRGINDNNTFYANSFTPIKAQVRLYPQ 299
QY 301 VCRHCTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRTLLNMDFTWEPKRAKLD 360
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 300 ICRHCTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRTLLNMDFTWEPKRAKLD 359
QY 361 KOGKVAMTSGHNDOSQWLQVDLLVPTKVGTGITOGAKDFGHVQVGSYKLAISNDGEHW 420
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 360 KOGKVAMTSGHNDOSQWLQVDLLVPTKVGTGITOGAKDFGHVQVGSYKLAISNDGEHW 419
QY 421 TVYODEKORKDVPOGNDNDTHRKKNVDPPIYARHILPLMSWYGRITLASELLGCCEE 480
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 420 MVHODEKORKDVPOGNDNDTHRKKNVDPPIYARHILPLMSWYGRITLASELLGCCEE 479
QY 481 E 481
DB 480 E 480

```

```

RESULT 4
US-08-659-235C-10
; Sequence 10, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quartermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.

```

```

; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659, 235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235C-10

```

```

Query Match          94.2%; Score 2519.5; DB 2; Length 480;
Best Local Similarity 94.2%; Pred. No. 1.2e-197;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

```

```

QY 1 MKRVAWLLVGLSLGVPOFGKGDICDPNCEGICLPLGLAVSFCSPGDTFDPNCS 60
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 1 MKHLVAMLLVGLSLGVPOFGKGDICDPNCEGICLPLGLAVSFCSPGDTFDPNCS 60
QY 61 SVVEVASDEEPTSGAGPCIPNCHNGGTCISEAYRGDTFGLGYCKCPRGNGIHCOHNI 120
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 61 SVVEVASDEEPTSGAGPCIPNCHNGGTCISEAYRGDTFGLGYCKCPRGNGIHCOHNI 120
QY 121 NECEVPECKNGGICITDLVANYSCGCEPGEFMRNCOYKCSGPLGIEGIIISNOQITASSSTH 180
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 121 NECEVPECKNGGICITDLVANYSCGCEPGEFMRNCOYKCSGPLGIEGIIISNOQITASSSTH 180
QY 181 RALFGLQKWYPYARLNKKGLINAMTAENDRMKRWIOLNORRMRYGVITOGAKRIGS 240
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 181 RALFGLQKWYPYARLNKKGLINAMTAENDRMW-IQINLORRMRYGVITOGAKRIGS 239
QY 241 PEYIKFYKIAVSNCKTAMKVKGTNEDMVRGINDNNTFYANSFTPIKAQVRLYPQ 300
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 240 PEYIKFYKIAVSNCKTAMKVKGTNEDMVRGINDNNTFYANSFTPIKAQVRLYPQ 299
QY 301 VCRHCTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRTLLNMDFTWEPKRAKLD 360
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 300 ICRHCTLRMELLCGELSGSEPLGKSGHIDYQITASSIFRTLLNMDFTWEPKRAKLD 359
QY 361 KOGKVAMTSGHNDOSQWLQVDLLVPTKVGTGITOGAKDFGHVQVGSYKLAISNDGEHW 420
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 360 KOGKVAMTSGHNDOSQWLQVDLLVPTKVGTGITOGAKDFGHVQVGSYKLAISNDGEHW 419
QY 421 TVYODEKORKDVPOGNDNDTHRKKNVDPPIYARHILPLMSWYGRITLASELLGCCEE 480
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 420 MVHODEKORKDVPOGNDNDTHRKKNVDPPIYARHILPLMSWYGRITLASELLGCCEE 479
QY 481 E 481

```

Db 480 E 480

RESULT 5
US-08-480-229C-21

; Sequence 21, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quartermou, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-480-229C-21

Query Match 64.2%; Score 1717; DB 2; Length 321;

Best Local Similarity 98.4%; Pred. No. 1.8e-132;

Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 158 CSGLPIEGGIISNOQITASTSTRALFGLQKWPYARLNKKGILINAWTAENDRMKRWI 217
|||||
Db 1 CSGLPIEGGIISNOQITASTSTRALFGLQKWPYARLNKKGILINAWTAENDRMKRWI 60
QY 218 QINLQKRMVTVITOGAKRIGSPYIKFYKIAYSDNCKTWMYKVGKTNEDMVEFRGND 277
|||||
Db 61 QINLQKRMVTVITOGAKRIGSPYIKFYKIAYSDNCKTWMYKVGKTNEDMVEFRGND 120
QY 278 NNTPYANSTPPIKAQYVRLYPVCRHCTLRMELLCGCELSGSEPLGKMSGHIDYQIT 337
|||||
Db 121 NNTPYANSTPPIKAQYVRLYPVCRHCTLRMELLCGCELSGSEPLGKMSGHIDYQIT 180
QY 338 ASSIFRTLNDMFTWEPKARLDKQGVNMTSGHNDOSQWLOVLLVPTKVGTITOGA 397
|||||
Db 181 ASSIFRTLNDMFTWEPKARLDKQGVNMTSGHNDOSQWLOVLLVPTKVGTITOGA 240
QY 398 KDFGHVQFVGSYKLAYSNDGEHMTVYODEKQKRDVFOGNEFDNDHRRKNVIDPPIYARHI 457
|||||
Db 241 KDFGHVQFVGSYKLAYSNDGEHMTVYODEKQKRDVFOGNEFDNDHRRKNVIDPPIYARHI 300

QY 458 RILPMSWYGRITLASELLGCT 478
|||||
Db 301 RILPMSWYGRITLASELLGCT 321RESULT 6
US-08-659-235C-21

; Sequence 21, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quartermou, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-659-235C-21

Query Match 64.2%; Score 1717; DB 2; Length 321;

Best Local Similarity 98.4%; Pred. No. 1.8e-132;

Matches 316; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 158 CSGLPIEGGIISNOQITASTSTRALFGLQKWPYARLNKKGILINAWTAENDRMKRWI 217
|||||
Db 1 CSGLPIEGGIISNOQITASTSTRALFGLQKWPYARLNKKGILINAWTAENDRMKRWI 60
QY 218 QINLQKRMVTVITOGAKRIGSPYIKFYKIAYSDNCKTWMYKVGKTNEDMVEFRGND 277
|||||
Db 61 QINLQKRMVTVITOGAKRIGSPYIKFYKIAYSDNCKTWMYKVGKTNEDMVEFRGND 120
QY 278 NNTPYANSTPPIKAQYVRLYPVCRHCTLRMELLCGCELSGSEPLGKMSGHIDYQIT 337
|||||
Db 121 NNTPYANSTPPIKAQYVRLYPVCRHCTLRMELLCGCELSGSEPLGKMSGHIDYQIT 180
QY 338 ASSIFRTLNDMFTWEPKARLDKQGVNMTSGHNDOSQWLOVLLVPTKVGTITOGA 397
|||||
Db 181 ASSIFRTLNDMFTWEPKARLDKQGVNMTSGHNDOSQWLOVLLVPTKVGTITOGA 240
QY 398 KDFGHVQFVGSYKLAYSNDGEHMTVYODEKQKRDVFOGNEFDNDHRRKNVIDPPIYARHI 457

Db 241 KDXHVPFVSSYKLAYSNDGEHMTVXODEKORDKYXGNFDMTHRKNIIDPPIYARH1 300
QY 458 RILPMSWGRTTLASELLGCT 478
Db 301 RILPMSWGRTTLASELLGCT 321

RESULT 7
US-08-162-402B-9
Sequence 9, Application US/08162402B
Patent No. 5972337
GENERAL INFORMATION:
APPLICANT: CERIANI, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
APPLICANT: LAROCCA, DAVID J.
TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
TITLE OF INVENTION: GLOBULE (HMEG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder & Poplawski
STREET: 444 South Flower St., 19th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-9

Query Match 42.8%; Score 1144; DB 2; Length 463;
Best Local Similarity 48.0%; Pred. No. 1.4e-85;
Matches 220; Conservative 72; Mismatches 144; Indels 22; Gaps 8;

QY 23 GDICDPNRCNGGICLPGGLAVGSFSCPCPGFTDPNCSVVEVASDBEETSGPCTPNP 82
Db 25 GDCPDSSLCINGGTCILTG-ODNDIYICLCPGFTGLVCNE-----TERGPCSPNP 72
QY 83 CHNGGTEIS-EAVRGDTFGYVCKPRGNGICHOHININECEVEPCCKNGICTDLVAVY 141
Db 73 CYNDAKCLVLDIORGDIIFTEYICQCPVGSIGHCETETNYMD---GEYMTTAVPNT 129
QY 142 SCECPGEF--MGRNCOYKSGPLGIEGGITISNOQTAASTHRLFGLOKWPYVYARLNR 199
Db 130 AVLPAPPTPLSNILASRCSQGLMEGAIADQSISASYVMGFMGLQRMGPFLARLYRT 189
QY 200 GLINAWTAANDRWKRNIQINLQKMKRVYGVITQGAKRIGSPETIFYKIAYSNDCKTWA 259

Db 190 GIVNAHNASNYSLP-WIQVNLIRKRVSGVMTQASRAGRAEYLKTFKVAYSILGRFFE 248
QY 260 MKYKGTNEDEWVRGINDNTPYANSFPPIKAOVVRLYPOVCRHCHLRMLLGCCELSG 319
Db 249 FTQDE-SGGDKERLGLNDDNNSLKVNFNLTAEYTRLYPSCHGCTLRRELLGCELHG 307
QY 320 CSEPLGKSGHIQDYQITFASSTFRTLMDMTWEPERKARLDKQGVNMTSGHNDOSQML 379
Db 308 CLEPLGLKNNITLPDSQMSASSYKTMNLRAPGVYHLGRDNOGKINMTAQSNARKEML 367
QY 380 QVDLVPKVTGIIITOGAKDEGHVOPVGSYKLAYSNDGEHMTVYODEKORDKYFQGNFD 439
Db 368 QVDLGFQROVYTGIIITOGARDFGHIOYVESYKVAHSDDDGVQMTVY--EFGSSKVFQGNLD 425
QY 440 NDIHRKNVTDPPITYARH1RILPMSWGRTTLASELLGCT 477
Db 426 NNSHKNIPEKPPMAKRVLPVSMHNRITTLLELLGC 463

RESULT 8
US-08-480-229C-29
Sequence 29, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-480-229C-29

Query Match 42.4%; Score 1133.5; DB 2; Length 221;
Best Local Similarity 90.0%; Pred. No. 4e-85;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 MKRSVAVMLVLSIGVQFGKGDICDPNRCNGGICLPGGLAVGSFSCPCDGTDPNCS 60

Db 1 MKHLVAMLLVGLSLGVPQFGKGDICPNPCENGICISGLADSPSCPCPEGAPNCS 60
QY 61 SYVEVASDEEPTISAGPCTPNPCNGNGTCETSEAYRGDTFTGYVCKCPRGFGNHCQNH 120
Db 61 SYVEVASDEEPTISAGPCTPNPCNGNGTCETSEAYRGDTFTGYVCKCPRGFGNHCQNH 120
QY 121 NECEVEPCCKNGICTDLVANYSCCEPGEFMRNCQYKCSGPLGIEGIIISNOOITASSNH 180
Db 121 NECEVEPCCKNGICTDLVANYSCCEPGEFMRNCQYKCSGPLGIEGIIISNOOITASSNH 180
QY 181 RALFGLQKWYPYARLNKKGILINMTAENDRMRWIOINL 221
Db 181 RALFGLQKWYPYARLNKKGILINMTAENDRMRWIOINL 221
RESULT 9
US-08-659-235C-29
; Sequence 29, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-08-659-235C-29
Query Match 42.4%; Score 1133.5; DB 2; Length 221;
Best Local Similarity 90.0%; Pred. No. 4e-85; Indels 1; Gaps 1;
Matches 199; Conservative 6; Mismatches 15;
QY 1 MKRSVAVMLLVGLSLGVPQFGKGDICDPNPCENGICLPLGLAVSFSCECPDGFDPNCS 60
Db 1 MKHLVAMLLVGLSLGVPQFGKGDICPNPCENGICISGLADSPSCPCPEGAPNCS 60
QY 61 SYVEVASDEEPTISAGPCTPNPCNGNGTCETSEAYRGDTFTGYVCKCPRGFGNHCQNH 120
Db 61 SYVEVASDEEPTISAGPCTPNPCNGNGTCETSEAYRGDTFTGYVCKCPRGFGNHCQNH 120

QY 121 NECEVEPCCKNGICTDLVANYSCCEPGEFMRNCQYKCSGPLGIEGIIISNOOITASSNH 180
Db 121 NECEVEPCCKNGICTDLVANYSCCEPGEFMRNCQYKCSGPLGIEGIIISNOOITASSNH 180
QY 181 RALFGLQKWYPYARLNKKGILINMTAENDRMRWIOINL 221
Db 181 RALFGLQKWYPYARLNKKGILINMTAENDRMRWIOINL 221
RESULT 10
US-08-162-402B-8
; Sequence 8, Application US/08162402B
; Patent No. 5972337
; GENERAL INFORMATION:
; APPLICANT: CERTANT, ROBERTO L.
; APPLICANT: PETERSON, JERRY A.
; APPLICANT: LAROCCA, DAVID J.
; TITLE OF INVENTION: 46 KALTON HUMAN MILK FAT
; TITLE OF INVENTION: GLOBULE (HMF) ANTIGEN, FRAGMENTS & FUSION PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder & Poplowski
; STREET: 444 South Flower St., 19th Floor
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,402B
; FILING DATE: 03-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Amzel, Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P66 38215
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-622-7700
; TELEFAX: 213-489-4210
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-162-402B-8
Query Match 41.9%; Score 1120.5; DB 2; Length 465;
Best Local Similarity 47.6%; Pred. No. 1.2e-83; Indels 31; Gaps 9;
Matches 225; Conservative 63; Mismatches 154;
QY 9 LLVGLSLGVPQFGKGDICDPNPCENGICLPLGLAVSFSCECPDGFDPNCSVVEVASD 68
Db 20 LLVGLD-----DPCDSSSLCGLNGTCLTG-QDNDIYCCLPREGFTGYVND- 62
QY 69 EEPPTISAGPCTPNPCNGNGTC-ETSEAYRGDTFTGYVCKCPRGFGNHCQNHNECEVEP 127
Db 63 ---TERGICSNPCPNHNGICLCEETISQEVRGDVPFSPYCTCTCKAGAHNCEFTYVYND- 117
QY 128 CKNGICTDLVANYSCCEPGEF--MGRNCQYKCSGPLGIEGIIISNOOITASSNHRLFLG 185
Db 118 -GEYMTTAVPNTAVPTPADPTDLISNMLASKCVPEPLGMENGINANSOIAASSVRYVFLG 175

QY 186 LQWVYVYARLNKKGLINAWTAENDRWKRWIOINLQKKRVGTITOGAKRIGSPYIK 245
176 LQHWVPELARLNRAAGVNAWTPSSND- NPMIOYNILRRMWVGVVYOGASRLASHETLK 234
QY 246 FYKAIYSNDGKTW-AMTKVKGTDNMYFRGNIDNNTPYANSFPPIKAQVRLYPOVCCR 304
Db 235 AFKAVYSLNGHEFPFIHDVYNNKHEFY--GNMKNNAVHVLFEFPVAQVRLYPTSCHT 292
QY 305 HCTLRMELGELSGSEPIGMSGHIQDQVITASSIFRTLNMDFTWEPKARLDQCK 364
Db 293 ACTLRPELGGELGNCANPLGLNNNSIPDKQITASSSYKTWGLHFSMNSYARLDKQGN 352
QY 365 VNANTSGHNDQSWLQVLDLVPRTKVTGIIITOGAKDFGHVOFVGSYKLAYSNDGEHMTVYQ 424
Db 353 FNWAVAGSYGNDQWLDQDLSKSEVTGIIITOGARNFGSVQFVASYKAYNSDANSMTYEQ 412
QY 425 DEKQRKRVQGNFNDNTHRKKNVDPPIYARHRIILPMSWYGRITLASSELLGC 477
Db 413 DPRGSSKIFPGMNDHSHKKNLFEPTILARYRIILFVAMHNRIALRELLGC 465

RESULT 11

US-08-162-402B-6
Sequence 6, Application US/08162402B

GENERAL INFORMATION:
APPLICANT: CERIANI, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
APPLICANT: LAROCCA, DAVID J.
TITLE OF INVENTION: 46 DALTON HUMAN MILK FAT
TITLE OF INVENTION: GLOBULE (HMEG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder & Poplowski
STREET: 444 South Flower St., 19th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-6

Query Match 38.1%; Score 1020.5; DB 2; Length 387;
Best Local Similarity 48.8%; Pred. No. 1.3e-75;
Matches 196; Conservative 55; Mismatches 108; Indels 43; Gaps 5;

QY 78 CTFPNCHNGTGC-EISEAYRGDTFTGYVCKCPRENGICQHININECEVEPCKNGICTD 136
Db 27 CSKNPCNHNGLEELSEEYRGDVFPSYTCCLKGYAGNH----- 65
QY 137 LVANYSCCEPGEPMGNCQYKCSGLPIEGGIISNOQITASSTHRALFGLOKWPYARL 196
Db 66 -----CETKCYEPLGMEGNINANSQIAASSVRYTFGLQHWPELARL 108
QY 197 NKKGLINAWTAENDRWKRWIOINLQKKRVGTITOGAKRIGSPYIKFYKAIYASNDGK 256
Db 109 NRGAVNNAWTPSSND- NPMIOYNILRRMWVGVVYOGASRLASHETLKAFKAVASLNGH 167
QY 257 TW-AMTKVKGTDNMYFRGNIDNNTPYANSFPPIKAQVRLYPOVCCRHCTLRHELLGC 315
Db 168 EPEFDHVNKKHKEFY--GNMKNNAVHVLFEFPVAQVRLYPTSCHTACTLRHELLGC 225
QY 316 ELSGCSEPIGMSGHIQDQVITASSIFRTLNMDFTWEPKARLDKQGVNANWTSGHNDQ 375
Db 226 ELNGCANPLGLNNNSIPDKQITASSSYKTWGLHFSMNSYARLDKQGNFANWAGSYGN 285
QY 376 SOMLOYDLVLPRTKVTGIIITOGAKDFGHVOFVGSYKLAYSNDGEHMTVYODEKQRKRVQ 435
Db 286 DQWLDQDLSKSEVTGIIITOGARNFGSVQFVASYKAYNSDANSMTYEQDPRGSSKIFP 345
QY 436 GNFDNDTHRKKNVDPPIYARHRIILPMSWYGRITLASSELLGC 477
Db 346 GNMNDHSHKKNLFEPTILARYRIILFVAMHNRIALRELLGC 387

RESULT 12

US-08-480-229C-20
Sequence 20, Application US/08480229C

GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALITY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-480-229C-20

Query Match	35.48;	Score 948;	DB 2;	Length 320;
-------------	--------	------------	-------	-------------

Matches 175; Conservative 50; Mismatches 95
 best local similarity 04.7%, freq. no. 0.06 100

150 ACCEPTECCTISNOCTAESCUPAEFCI QKWBVYABI NKKCI TNAWTA AENDPBRBWT 317

[illegible]

RESULT 13
US-08-659-235C-20

```

;      TOPOLOGY:  unknown
;      MOLECULE TYPE:  protein
US-08-659-235C-20

```

Query Match 35.48; Score 948; DB 2; Length 320

Best Local Similarity	54.7%;	Pred. No.	8.5e-70;						
Matches	175:	Conservative	50:	Mismatches	95:	Indels	0:	Gaps	0:

THE UNIVERSITY OF MICHIGAN LIBRARY

[illegible]

```

1 RESULT 14
2 US-08-746-111-5
3 ; Sequence 5, Application US/08746111
4 ; Patent No. 6086778
5 ;
6 ; GENERAL INFORMATION:
7 ; APPLICANT: Ginsburg, David
8 ; APPLICANT: Cui, Jisong
9 ; TITLE OF INVENTION: Compositions And Methods For Screening
10 ; TITLE OF INVENTION: Compounds For Anticoagulant Activity
11 ; NUMBER OF SEQUENCES: 54
12 ;
13 ; CORRESPONDENCE ADDRESS:
14 ; ADDRESSEE: Medlen & Carroll, LLP
15 ; STREET: 220 Montgomery Street, Suite 2200
16 ; CITY: San Francisco
17 ; STATE: California
18 ; COUNTRY: United States of America
19 ; ZIP: 94104
20 ;
21 ; COMPUTER READABLE FORM:
22 ; MEDIUM TYPE: Floppy disk
23 ; COMPUTER: IBM PC compatible
24 ; OPERATING SYSTEM: PC-DOS/MS-DOS
25 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
26 ; CURRENT APPLICATION DATA:
27 ; APPLICATION NUMBER: US/08/746,111
28 ; FILING DATE: 06-NOV-1996
29 ;
30 ; CLASSIFICATION:
31 ; ATTORNEY/AGENT INFORMATION:
32 ; NAME: Ingolia, Diane E.
33 ; REGISTRATION NUMBER: 40,027
34 ; REFERENCE/DOCKET NUMBER: UM-02536
35 ; TELECOMMUNICATION INFORMATION:
36 ; TELEPHONE: (415) 705-8410
37 ; TELEFAX: (415) 397-8338
38 ;
39 ; INFORMATION FOR SEQ ID NO: 5:
40 ; SEQUENCE CHARACTERISTICS:
41 ; LENGTH: 2183 amino acids
42 ; TYPE: amino acid
43 ; TOPOLOGY: linear
44 ;
45 ; MOLECULE TYPE: protein
46 ;

```


Db 310 WLIQDLSQRNVYGIITQCARDFGHIQYAAKVAASDQVSWTEYRDQALGKIFPEN 369

QY 437 FDNDRHRNVIDPPIYARFIRLPSWYGRITLRSLLGC 476

Db 370 IDNNSHKKNMPEPFLRFVRIPLVAMHNRITLRLVLLGC 409

RESULT 2

A36479
milk fat globule membrane protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A36479
R:Studivs, J.D.; Lekutis, C.; Singer, K.L.; Bui, A.; Yuzuki, D.; Srinivasan, U.; Parry, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990
A:Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the ex
A:Reference number: A36479; MUID:91046008
A:Accession: A36479
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-463 <STU>
A:Cross-references: GB:M38337; NID:9199142; PIDN:AAA9534.1; PID:9199143
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C:Keywords: membrane protein
F:28-60/Domain: EGF homology <EG1>
F:68-107/Domain: EGF homology <EG2>
F:147-303/Domain: discoidin I amino-terminal homology <DN1>
F:307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match

Best Local Similarity 43.8%; Score 1168.5; DB 1; Length 463;
Matches 223; Conservative 74; Mismatches 138; Indels 23; Gaps 8;

QY 23 GDICNPNCENGICLSGLADDSFSCBCEBGFAGPNCSSVVEYASDEKPTSAQPCIPNP 82

Db 25 GDFCDSSICLNGGTCILG-QDNDIYCLCEPGETGLVNE-----TERGCSFNP 72

QY 83 CHNGSTCEIS-EAYRGDTFIVGCKPRGFNGIHCHHINECEAPCRNGICTDLVANY 141

Db 73 CYNMAKCLVTDYORGDIFTEYICQCPVGSIGIHETETNYND---GTYMTTAVPNT 129

QY 142 SCECPGEF--MGRNCOYKSGHLIEGGLISMOQITASSNHRALFGLQWYVYARLANK 199

Db 130 AAPPFAPPPDLSNNLASRSTGLGMEGALADSOISASVYMGFMQWRGPELALYFT 189

QY 200 GLINAMTAENDRMPIQIINLQKRMVTVITQAKRIGSEPEYIKSYKLIASNDGKTWAM 259

Db 190 GIVNWMHANSYDSKPMIOVNLKRMKRVSGVMQGSASRAGREYLKTPKVAVSLDGKREF 249

QY 260 YK-VKGTNEEMVPRGVNDNTPYANSFPPRIKAOYVRLYPOICRRHCTLRMLLGCESG 318

Db 250 IODESGGKKE--FLGNLDNNSLKNMNFNTLEAOYIRLYPSCHRGCTLRFLGCELHG 307

QY 319 CSEPLGKMSGHIDYQITASVFRILNMDFTPEPRKARLDKQKYNAMTSGHNDOSWL 378

Db 308 CLEPGLKNNITIPDSOMASASSYKWNLRAGFWPHLGRDLNCKINAMTAOSNKAEML 367

QY 379 QVDLVPTKVGITITQAKDFGHVQFVSGYKLIASNDGEHMYHODEKORKKDFVEQNF 438

Db 368 QVDLCTQPOVGTITQCARDFGHIVYVESYKVAHSDGVQMTYV--EEQSSVFEQGLND 425

QY 439 NDTHKKNVIDPPIYARFIRLPSWYGRITLRSLLGC 476

Db 426 NNSHKKNIFEKPFMARIVRVLPSWHNRITLRLVLLGC 463

RESULT 3

JC4915
ags protein precursor - rat
N:Alternate names: O-acetyl-Cd3 ganglioside
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 20-Jun-2000

C:Accession: JC4915
R:Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y.
Biochem. Biophys. Res. Commun. 225, 932-938, 1996
A:Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.
A:Reference number: JC4915; MUID:96374422
A:Accession: JC4915
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-427 <OGU>
A:Cross-references: DDBJ:D84068; NID:91620006; PIDN:BAI12210.1; PID:91620007
A:Experimental source: CSR cell
C:Comment: This protein is required for the O-acetylation of distalganglioside siall
C:Genetics:

A:Gene: ags
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF hom
F:1-21/Domain: signal sequence #status predicted <SIG>
F:28-60/Domain: EGF homology <EG1>
F:68-107/Domain: EGF homology <EG2>
F:110-267/Domain: discoidin I amino-terminal homology <DN1>
F:271-427/Domain: discoidin I amino-terminal homology <DN2>

Query Match

Best Local Similarity 43.7%; Score 1166.5; DB 2; Length 427;
Matches 220; Conservative 65; Mismatches 117; Indels 53; Gaps 6;

QY 23 GDICNPNCENGICLSGLADDSFSCBCEBGFAGPNCSSVVEYASDEKPTSAQPCIPNP 82

Db 25 GDFCDSSICLNGGTCILG-QDNDIYCLCEPGETGLVNE-----TERGCSFNP 72

QY 83 CHNGSTCEIS-EAYRGDTFIVGCKPRGFNGIHCHHINECEAPCRNGICTDLVANY 141

Db 73 CFHMAKCLVTDYORGDIFTEYICQCPVGSIGIHETETNYND---GTYMTTAVPNT 108

QY 142 SCECPGEF--MGRNCOYKSGHLIEGGLISMOQITASSNHRALFGLQWYVYARLANK 201

Db 109 -----LG-----CSTKGLGEGALADSOISASVYMGFMQWRGPELALYRGI 154

QY 202 INAMTAENDRMPIQIINLQKRMVTVITQAKRIGSEPEYIKSYKLIASNDGKTWAM 261

Db 155 VNAMTAENDRMPIQIINLQKRMVTVITQAKRIGSEPEYIKSYKLIASNDGKTWAM 214

QY 262 VKGTNEEMVPRGVNDNTPYANSFPPRIKAOYVRLYPOICRRHCTLRMLLGCESG 321

Db 215 DESGTGKREMGHNDNLSKIMFNPNTLEAOYIRLYPSCHRGCTLRFLGCELHGSE 274

QY 322 PLGKMSGHIDYQITASVFRILNMDFTPEPRKARLDKQKYNAMTSGHNDOSWLQVD 381

Db 275 PLGKNNITIPDSOMASASSYKWNLRAGFWPHLGRDLNCKINAMTAOSNKAEMLQVD 334

QY 382 LLVFTKYTGITITQAKDFGHVQFVSGYKLIASNDGEHMYHODEKORKKDFVEQNF 441

Db 335 LGTKKATYTGITITQAKDFGHVQFVSGYKLIASNDGEHMYHODEKORKKDFVEQNF 392

QY 442 HRKNVIDPPIYARFIRLPSWYGRITLRSLLGC 476

Db 393 HKKNIFEKPFMARIVRVLPSWHNRITLRLVLLGC 427

RESULT 4

PAS-6/7 protein precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
C:Accession: S74211; S78114; S24181
R:Harregard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.
Eur. J. Biochem. 240, 628-636, 1996
A:Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat g
A:Reference number: S74211; MUID:97008954
A:Accession: S74211
A:Molecule type: mRNA
A:Residues: 1-427 <HVA>
A:Cross-references: EMBL:X91895; NID:91632778; PIDN:CAA62997.1; PID:91632779

A:Residues: 1-2183 <YAN>
A:Cross-references: EMBL:052925; NID:g3219690; PID:g3219691; PIDN:AAC9553.1
C:Function:
A:Pathway: blood coagulation
C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F:350/682/Domain: ferroxidase repeat homology <FOX1>
F:1341-1864/Domain: ferroxidase repeat homology <FOX2>

Query Match 26.0%; Score 694; DB 2; Length 2183;
Best Local Similarity 43.8%; Pred. No. 2.9e-42;
Matches 144; Conservative 51; Mismatches 112; Indels 22; Gaps 5;

QY 157 KCSGHLGEGGIIISNOQITASSNHRALFGLOKWPYARLNKGLIAM-----TAENR 212
DB 1865 ECKMKSLSTGVISDSQIRASE-----LYWEPLRLAKNNASYNAMSTKLTALDPPI 1918
QY 213 WPTQIOLNRRKRVYVITOGAKRIGSPYIKS-----YRIASNDGKTAMVKVKTNE 267
DB 1919 KPWIOVMQKEVVVVTGIOTOGAK-----HYLKSCTTEFFOVAYSSDQTNMOLFRKSGKS 1973
QY 268 EMVFRGNVDNTPYANSFTPIKAOYVRLYPOICRHCTLMELGCGELSGCEPLGMS 327
DB 1974 VMYFGNSDGSITKENRLDPPIVARIHPTKSYNRPITLLELQCEVNGCSTPLGLD 2033
QY 328 GHIDYDITASSVFRLLNMDFTWEPKRAKLDKQKNAWTSQHNDQSCWLOVDLLVPRK 387
DB 2034 GRIDQKLTASSFRKSKWGDY--WPSLARLVAQGRVAMQAKNNKQWLOVDLKLTK 2091
QY 388 VTGIIITOGANDFGHVOFGVSYSKLAVSNDGEHMMVHODEKORKDVFQGNFNDTHKRVY 447
DB 2092 VTAIVTQGCSSLSEMYVKSYSLOYSDQVAMKPYRQKSSWDKIEFGNSNTGKMKNF 2151
QY 448 DPTIARFIRLLPMSWYGRITLRESLGC 476
DB 2152 NPILSRFIRLIPKTNQSLALRELFEGC 2180

RESULT 7
KRFH05
coagulation factor V precursor [validated] - human
N:Alternate names: coagulation labile factor; proaccelerin
C:Species: Homo sapiens (man)
C:Date: 19-May-1989 #sequence, revision 02-Jun-1995 #text_change 08-Dec-2000
C:Accession: A56172; A42344; A28028; A27498; A25897
R:Cripe, L.D.; Moore, K.D.; Kane, W.H.
Biochemistry 31, 3777-3785, 1992
A:Title: Structure of the gene for human coagulation factor V.
A:Reference number: A42344; MOID:92232686
A:Accession: A56172
A:Molecule type: DNA
A:Residues: 1-2224 <CRI>
A:Cross-references: GB:J05368
A:Accession: A42344
A:Molecule type: DNA
A:Residues: 48-58;79-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-542;
2070;2111-2120;2172-2181 <CRO>
R:Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kaufm
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
A:Title: Complete cDNA and derived amino acid sequence of human factor V.
A:Reference number: A28028; MOID:8726086
A:Accession: A28028
A:Molecule type: mRNA
A:Residues: 1-857, 'R', 859-864, 'R', 866-924, 'E', 926-1763, 'I', 1765-2212, 'T', 2214-2224 <JEN>
A:Cross-references: GB:M16967
A:Note: Parts of this sequence, including the amino end of the mature protein, were dete
Biochemistry 26, 6508-6514, 1987
R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
A:Title: Cloning of cDNA coding for the heavy chain region and connecting region of hum
A:Reference number: A27498; MOID:88107560
A:Accession: A27498
A:Molecule type: mRNA

A:Residues: 1-1284, 'I', 1286-1600 <KAN>
A:Cross-references: GB:M17785
A:Note: parts of this sequence were determined by protein sequencing
R:Kane, W.H.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986
A:Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homo
A:Reference number: A25897; MOID:86313655
A:Accession: A25897
A:Molecule type: mRNA
A:Residues: 1188-1215, 1315-2224 <KA2>
A:Cross-references: GB:M14335
A:Note: parts of this sequence were determined by protein sequencing
R:Keller, F.G.; Otey, T.L.; Quinn-Allen, M.A.; Kane, W.H.
Biochemistry 34, 4118-4124, 1995
A:Title: Thrombin catalyzed activation of recombinant human factor V.
A:Reference number: A56139; MOID:95210278
A:Contents: annotation; thrombin cleavage sites
C:Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C:Genetics:
A:Gene: GDB:F5
A:Cross-references: GDB:119896; OMIM:227400
A:Map position: 1q23-1q23
A:Introns: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1;
C:Function:
A:Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa p
A:Pathway: blood coagulation
C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plas
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-2224/Product: coagulation factor V #status predicted <MAT>
F:29-737/Product: coagulation factor Va heavy chain #status experimental <VAH>
F:29-345/Domain: A1 <DA1>
F:33-329/Domain: ferroxidase repeat homology <FO1>
F:346-691/Domain: A2 <DA2>
F:351-684/Domain: ferroxidase repeat homology <FO2>
F:692-1573/Domain: B <DOB>
F:1183-1461/Region: 9-residue repeats (O-X-T/N-L-S-P-D-L-S)
F:1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>
F:1574-1905/Domain: A3 <DA3>
F:1581-1905/Domain: ferroxidase repeat homology <FO3>
F:1667-1765/Region: phospholipid binding #status predicted
F:1906-2064/Domain: C1 <DC1>
F:1906-2061/Domain: discoidin I amino-terminal homology <DNI>
F:2065-2224/Domain: C2 <DC2>
F:2065-2221/Domain: discoidin I amino-terminal homology <DN2>
F:51-55;239;297;460;468;554;741;752;760;776;782;821;938;977;1074;1083;1103;1106;1479;
F:167-193;248-329;500-526;603-684;1725-1751;1907-2061;2086-2221/Disulfide bonds: #sta
F:334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F:363;693;1546/Binding site: Arg-Asn (protein C) #status predicted
F:376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F:382;1338/Binding site: carboxylate (Asn) (covalent) #status absent
F:534-535/Cleavage site: Arg-Gly (protein C) #status predicted
F:737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
F:1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experime
F:1573-1574/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 25.6%; Score 684; DB 1; Length 2224;
Best Local Similarity 42.4%; Pred. No. 1.6e-41;
Matches 143; Conservative 55; Mismatches 123; Indels 16; Gaps 4;

QY 144 ECPPEFMRGRNCKYKSGHGLGEGIIISNOQITASSNHRALFGLOKWPYARLNKGLIN 203
DB 1897 QTPFLINDRCRMP-----MGLSTGIISDSQIRASE-----FLGYWEPLRLNNGGSVN 1946
QY 204 AWT-----AAENDWPIQIOLNRRKRVYVITOGAKRIGSPYIKSYRIASNDGKTAM 259
DB 1947 AWSVEKLAAPFASKPWIOVMQKEVITITGOTOGAKRILKSCFTTEFYVASSNOIMDI 2006
QY 260 YKVGKINEEWFGRGNVDNTPYANSFTPIKAOYVRLYPOICRHCTLMELGCGELSGC 319
DB 2007 FKGNSTRNVMYFNGNSDASTIKENQFDPPIVARIHPTKSYNRPITLLELQCEVNGC 2066

OY 370 GHNDOSCHLQYVDLLVPTKVTGIIITOGAKDFGHVOFVGSYKLAYSNDGEHMMVHODEKOR 429
Db 1154 LYNTFG-WVMEDEFGPRNITGILIKGND----GWTISYKVLVTSDEFTFNPAID-KDGK 1207
OY 430 DKVFOGNFNDNDTHRKNVLDPPPIYARFIRILPWSWYGRITLRSELLGCAE 478
Db 1208 EKIFPANFDGIVSVTNEFHPPIRARVTLKVLPOKWNKNIELRIEPIGCFE 1256

Search completed: May 23, 2001, 06:12:39
Job time: 73 sec

FT DISULFID 252 409 BY SIMILARITY.
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 409 AA: 45725 MW: B0C07AF80029277A CRC64:

Query Match 44.4%; Score 1184.5; DB 1; Length 409;
 Best Local Similarity 48.7%; Pred. No. 116-82;
 Matches 224; Conservative 65; Mismatches 112; Indels 59; Gaps 7;

QY 23 GDICNPNCENGICISGLAD---SFSCCECEGFCAGPNCSSVVEASDEKPTSGPCIT 79
 DB 3 GDFCDSSICLNGGTCCL-LDQDPQKPFHCLCEGFTGLICNE-----TEGPCSP 49
 QY 80 PPHCHNGSTCE-ISEAYRGDTFIGYCKCPRGFNIRHONHINCEAEPCRNNGICTDLV 138
 DB 50 PPHCHNDAECVDDAHRGVDFEYCYCKCPHGTGTH----- 86
 QY 139 ANYSCCEPGEFPMRNCQYKCSCHLIEGGIISNOQITASSNHRALFGLQKWPYVARLTK 198
 DB 87 -----CEIICNAPLGMETGAIADFQISASSMHLGFMGLQRMWAEELARLHR 131
 QY 199 KGLINAMTAENDRMPIQIINTLQKRRVGVITQCAKRGISEPEYIKAYKIAVNDGKTMA 258
 DB 132 AGIVAMATASNYDRNPWQIYNLRRKRYGVITQCASAGSAEAKTKFYAYSTDGRKQ 191
 QY 259 MKKAKYINE--EMVFRGNVDNNTPVANSTPPKAKOYRLYPQICRRHCTLRMELLGCEL 316
 DB 192 F--IQGAESGDKLFEMGNLDSGLKYNLEFVEQVYRALVPIICHHGCTLRRELLGCEL 249
 QY 317 SCSEPEPLGKSGHIDQYQITASSVFETLMMDFTWEPKRALDKQGVNATYSGHNDQSO 376
 DB 250 SCACAPLGLKONTIPNKQITASSFYRTWGLSAFMSVYFYARLDNQGFNMTQMSAAE 309
 QY 377 WLOYDLVPRKVGITITQAGKDGHVQFGSKYKLAASDGHMAYHODEQRDKYFOCN 436
 DB 310 WLOIDLSQRRVYITIIQAGRDGHQIYAAKVAISDGVSMTEPYRDQALGKIFPEN 369
 QY 437 FDNDRHKNVDPPIYARFTRILPMSWYGRITLRSSELLGC 476
 DB 370 LDNNSHKKNMETPELRFVRIPLVAMHNRITLRVELLGC 409

RESULT 2
 MFGM_MOUSE ID MFGM_MOUSE STANDARD; PRT: 463 AA.

AC 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAV GLOBULE-EGF FACTOR 8) (MFG-EB) (MFGM)
 DE (SPERM SURFACE PROTEIN SP47) (MP47).
 GN MFGEB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-35.
 RC TISSUE=Mammary gland;
 RX MEDLINE=91046008; PubMed=2122462;
 RA Stubbs J.D., Lekutis C., Singer K.L., Bul A., Yuzuki D.,
 RA Srinivasan U., Parry G.;
 RT "CDNA cloning of a mouse mammary epithelial cell surface protein
 RT to factor VIII-like sequences."
 RT Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
 RN [2]
 RP SEQUENCE OF 23-463 FROM N.A.
 RC TISSUE=Testis;
 RA Enslin M.A.;
 RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-

CC BINDING PROTEIN.
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND
 CC SPERMATOZOAN.
 CC -1- DEVELOPMENTAL STAGE: MRNA EXPRESSION IS DETECTABLE IN MAMMARY
 CC TISSUE FROM NONPREGNANT ANIMALS & MAXIMAL IN THE LACTATING GLAND.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 P5/8 TYPE C DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M38337; AAA39534.1; -
 CC EMBL: Y11684; CAA72380.1; -
 CC PIR: A36479; A36479.
 CC HSP: P00740; 11XA.
 CC MGD: MGI:102768; Mfge8.
 CC InterPro: IPR000421;
 CC InterPro: IPR000561; -
 CC InterPro: IPR001436; -
 CC Pfam: PF00754; F5_P8_Type_C; 2.
 CC PRINTS: PR00010; EGFBLD.
 CC PROSITE: PS00022; EGF_1; 2.
 CC PROSITE: PS0186; EGF_2; 2.
 CC PROSITE: PS01285; FAS8C_1; 2.
 CC PROSITE: PS01286; FAS8C_2; 2.
 CC KW Signal. Glycoprotein. Repeat: EGF-like domain; Milk.
 FT SIGNAL 1 22
 FT CHAIN 23 463 LACTADHERIN.
 FT DOMAIN 24 61 EGF-LIKE 1.
 FT DOMAIN 64 108 EGF-LIKE 2.
 FT DOMAIN 148 303 F5/8 TYPE C 1.
 FT DOMAIN 308 463 F5/8 TYPE C 2.
 FT SITE 87 89 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 28 39 BY SIMILARITY.
 FT DISULFID 33 49 BY SIMILARITY.
 FT DISULFID 51 60 BY SIMILARITY.
 FT DISULFID 68 79 BY SIMILARITY.
 FT DISULFID 73 96 BY SIMILARITY.
 FT DISULFID 98 107 BY SIMILARITY.
 FT DISULFID 148 303 BY SIMILARITY.
 FT DISULFID 290 294 BY SIMILARITY.
 FT DISULFID 308 463 BY SIMILARITY.
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 30 30 S -> F (IN REF. 2).
 FT CONFLICT 35 35 N -> D (IN REF. 1; AA SEQUENCE)
 FT CONFLICT 110 147 EMTYNDGEYFTFAVTPAFTPAFTDLSNNLASR ->
 FT CONFLICT 168 168 G (IN REF. 2).
 FT CONFLICT 196 196 Y -> S (IN REF. 2).
 FT CONFLICT 309 309 H -> T (IN REF. 2).
 FT CONFLICT 395 395 L -> S (IN REF. 2).
 FT CONFLICT 395 395 E -> A (IN REF. 2).
 SO SEQUENCE 463 AA: 51465 MW: D7B8C6CEBBA72AD CRC64:
 Query Match 43.8%; Score 1168.5; DB 1; Length 463;
 Best Local Similarity 48.7%; Pred. No. 2e-61;
 Matches 223; Conservative 74; Mismatches 138; Indels 23; Gaps 8;
 QY 23 GDICNPNCENGICISGLADSFSCCECEGFCAGPNCSSVVEASDEKPTSGPCIPNP 82
 DB 25 GDFCDSSICLNGGTCCLG-ODNDIYCLCEPCEGTGLVNE-----TERGPCSPNP 72
 QY 83 CHNGSTCEIS-EAYRGDTFIGYCKCPRGFNIRHONHINCEAEPCRNNGICTDLVANY 141

```

Db 73 CYNDAKLVLPDQROGDIPEYICQCPVSGIHCEETENYNDL--GEYMTTAVPNT 129
Oy 142 SCEPGEF--MGNCOYKCSGHGIGTGGIISNOQTASSNHRALFGLQKVPYARLNK 139
Db 130 AVPTAPPTPDLSSNLSRSTQLSMEGGAIDQSLASIVYMGVGRKQRPETLARLYKT 189
Oy 200 GLINAMTAENDRWPIQIINLQRMKRVTVITQAKRIGSEPIKSKYIAVSDKTKWAM 239
Db 190 GIVAMAMASVYDSKPMIOWLLKMKRVSGVMQGASRAGRAEYKTEKFAVSLDGRFEF 249
Oy 260 YK-VKGTNEEMVFRGVNDNTFPYNSFPPIKAQVVRILYPOICRRHCTLRMELLCGLSG 318
Db 250 IQDSSGDKK--FLGNIDNNLSLKVNFNTLEAQYIRLTPVSCHRGCTLRRELLGCELDHG 307
Oy 319 CSEPLGKSGHIDQYQITASSVFRILNMDFTWEPKRAIDKQGVNANMISGHNDOSQWL 378
Db 308 CLEPLGKNTITPDSOMASASSYKTNWNLKAFGMYEHLRLDQKINAMTAQNSAKEWL 367
Oy 379 QVDELVPKTVGIIITOGAKDFGVQVGSYKTLAYSNDEHMMVHODEKORKDYQGNPD 438
Db 368 QVDELGTQRYTGIIITOGARDFGHIQYVESYKVAHSDQGVQMTVY--EEOGSSKVFQGNLD 425
Oy 439 NDRHRKAVIDPIYARIRILPMSWYGRITLSEILGC 476
Db 426 NNSHKNIIFKRPKARVYRVLPSWNNRITLRLLELLGC 463

```

```

RESULT 3
MSGM_RAT STANDARD: PRT: 427 AA.
ID MSGM_RAT STANDARD: PRT: 427 AA.
AC P70490:
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-BGF FACTOR 8) (MFG-E8) (O-
GN ACETYL GD3 GANGLIOSIDE SYNTHASE) (AGS) (MFGM).
GN MFGE8 OR AGS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OY NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96374422; PubMed=8780713;
RA Ogura K., Nara K., Watanabe Y., Kohno K., Tai T., Sanai Y.;
RT "Cloning and expression of cDNA for O-acetylation of GD3
RT ganglioside."
RT Biochem. Biophys. Res. Commun. 225:932-938(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. SEEMS TO
CC PARTICIPATE IN THE O-ACETYLATION OF GD3 GANGLIOSIDE SIALIC ACID.
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: SPLEEN, LUNG, HEART, BRAIN AND MUSCLE.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D84068; BAA12210.1;
CC HSSP: P00740; 1IXA.
CC InterPro: IPR000421;
CC InterPro: IPR000561;
CC InterPro: IPR001438;
CC Pfam: PF00008; EGF_2.
CC Pfam: PF00754; F5_F8_Type_C_2.
CC PRINTS: PR00010; EGFBL00D.

```

```

DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FAS8C_1; 2.
DR PROSITE: PS01286; FAS8C_2; 2.
KM Signal: Glycoprotein Repeat; EGF-like domain; Milk.
FT SIGNAL 1 22
FT CHAIN 23 427
FT DOMAIN 23 427
FT DOMAIN 64 108
FT DOMAIN 111 267
FT DOMAIN 272 427
FT DISULFID 28 39
FT DISULFID 33 49
FT DISULFID 51 60
FT DISULFID 68 79
FT DISULFID 73 96
FT DISULFID 98 107
FT DISULFID 111 267
FT DISULFID 254 258
FT DISULFID 272 427
FT SITE 87 89
FT CARBOHYD 61 61
FT CARBOHYD 230 230
FT CARBOHYD 280 280
FT CARBOHYD 390 390
SQ SEQUENCE 427 AA; 47413 MM; EABC8631F3EE6047 CRC64;

```

```

Query Match 43.7%; Score 116.5; DB 1; Length 427;
Best Local Similarity 48.4%; Pred. No. 2.6e-81;
Matches 220; Conservative 65; Mismatches 117; Indels 53; Gaps 6;

```

```

Oy 23 GDICNPNCENGICISGLADSFSCPEGAGNCSSYVEVASDEKPTASAGPCIPNP 82
Db 25 GPCDSSCLNGTCLMG QDNIDITLCPGFTGLVNC-----TEKPCSPNP 72
Oy 83 CHNGGTGEISE-AVRGDTFYGVCCKPRFNGICHOHNIINCEAEPCRNIGICTDLYANT 141
Db 73 CFHDAKLVIEDIQRGDIFPEYICQCPVSGIHCE----- 108
Oy 142 SCEPGEFMRNCQYKCSGHGIGTGGIISNOQTASSNHRALFGLQKVPYARLNKGL 201
Db 109 -----LG-----CSTLGLGEGALADSOIASSVYMGVGRKQRPETLARLYGTGI 154
Oy 202 INAMTAENDRWPIQIINLQRMKRVTVITQAKRIGSEPIKSKYIAVSDKTKWAMK 261
Db 155 VNAMTASSYSKPMIOWLLKMKRVSGVMQGASRAGRAEYKTEKFAVSLDGRFEFQ 214
Oy 262 VKGTNEEMVFRGVNDNTFPYNSFPPIKAQVVRILYPOICRRHCTLRMELLCGLSCSE 321
Db 215 DESGTGDKFEGMGNDDNNSKTNMFNTLEAQYIRLTPVSCHRGCTLRRELLGCELDHSCSE 274
Oy 322 PLGKSGHIDQYQITASSVFRILNMDFTWEPKRAIDKQGVNANMISGHNDOSQWLQYD 381
Db 275 PLGLKNTITPDSOMASASSYKTNWNLKAFGMYEHLRLDQKINAMTAQNSAKEWLQYD 334
Oy 382 LLVPRKTVGIIITOGAKDFGVQVGSYKTLAYSNDEHMMVHODEKORKDYQGNPDNT 441
Db 335 LGTKKTVGIIITOGARDFGHIQYVESYKVAHSDQGVQMTVY--EEOGSSKVFQGNLDNNS 392
Oy 442 HRKNVVIDPIYARIRILPMSWYGRITLSEILGC 476
Db 393 HKNIIFKRPKARVYRVLPSWNNRITLRLLELLGC 427

```

```

RESULT 4
MSGM_BOVIN STANDARD: PRT: 427 AA.
ID MSGM_BOVIN STANDARD: PRT: 427 AA.
AC Q95114; O27959; P79344;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-BGF FACTOR 8) (MFG-E8)

```

(MGp57/53) (PAS-6/PAS-7 GLYCOPROTEIN) (MFGM) (SPERM SURFACE PROTEIN SP47) (BP47) (COMPONENTS 15/16).

DE MGp57
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RC SPFAIR-HOLSTEIN; TISSUE=Mammary gland;
RX MEDLINE=97008954; PubMed=8856064;
RA Hvarregard J., Andersen M.H., Berglund L., Rasmussen J.T.,
RA Petersen T.E.,
RT "Characterization of glycoprotein PAS-6/7 from membranes of bovine
RT milk fat globules.";
RL Eur. J. Biochem. 240:628-636(1996).
RN
RP SEQUENCE OF 18-427 FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=96125736; PubMed=8541316;
RA Aoki N., Kishi M., Taniguchi Y., Adachi T., Nakamura R.,
RA Matsuda T.,
RT "Molecular cloning of glycoprotein antigens MGp57/53 recognized by
RT monoclonal antibodies raised against bovine milk fat globule
RL Biochim. Biophys. Acta 1245:385-391(1995).
RN
RP SEQUENCE OF 19-427 FROM N.A.
RC TISSUE=Testis;
RA Esselin M.A.,
RN Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
RP
RP SEQUENCE OF 140-146; 174-187; 233-246 AND 422-427.
RC TISSUE=Milk;
RX MEDLINE=93250576; PubMed=8485470;
RA Maier I.H., Banghart L.R., Lane W.S.,
RT "The major fat-globule membrane proteins, bovine components 15/16 and
RT guinea-pig gp 55, are homologous to MGF-E8, a murine glycoprotein
RT containing epidermal growth factor-like and factor V/VIII-like
RT sequences.";
RL Biochem. Mol. Biol. Int. 29:545-554(1993).
CC
CC -1- FUNCTION: PROBABLY ASSOCIATES WITH PHOSPHOLIPIDS ON THE SURFACE OF
CC MAMMARY EPITHELIAL CELLS AND MILK FAT GLOBULES. ZONA PELLUCIDA-
CC BINDING PROTEIN.
CC
CC -1- ALTERNATIVE PRODUCTS: 2 ISOPFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM
CC LACKS 53 AMINO ACIDS WITHIN THE F5/8 TYPE C 1 DOMAIN.
CC
CC -1- TISSUE SPECIFICITY: MILK AND SPERMATOZOAN.
CC
CC -1- PROBABLY FUC AS REDUCING TERMINAL SUGAR.
CC
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sdb.ch).
CC
CC -----
DR EMBL: X91895; CNA62997.1;
DR EMBL: S80643; AAB35894.2;
DR EMBL: Y11719; CAA72406.1;
DR HSSP: P00740; IIXA;
DR InterPro: IPR000421;
DR InterPro: IPR000561;
DR Pfam: PR00008; EGF_2;
DR Pfam: PF00734; F5_F8_type_C_2;
DR PROSITE: PS00022; EGF_1; 2;
DR PROSITE: PS01186; EGF_2; 2;
DR PROSITE: PS01285; FA58C_1; 2.

DR PROSITE: PS01286; FA58C_2; 2.
KW Signal: Glycoprotein; Milk; Repeat; EGF-like domain;
KW Alternative splicing
FT SIGNAL 1 18
FT CHAIN 19 427
FT DOMAIN 20 59
FT DOMAIN 62 109
FT DOMAIN 109 265
FT DOMAIN 270 427
FT SITE 85 87
FT DISULFID 24 35
FT DISULFID 29 47
FT DISULFID 49 58
FT DISULFID 66 77
FT DISULFID 71 94
FT DISULFID 96 105
FT DISULFID 109 265
FT DISULFID 252 256
FT DISULFID 270 427
FT CARBOHYD 27 27
FT CARBOHYD 34 34
FT CARBOHYD 59 59
FT CARBOHYD 227 227
FT CARBOHYD 169 221
FT VARSPLIC 19 19
FT CONFLICT 28 28
FT CONFLICT 19 28
SQ SEQUENCE 427 AA; 47411 MW; 4CBEBE3A1DC4EB24 CRC64;
O-LINKED (FUC. . .) (IN PAS-6).
O-LINKED (FUC. . .) (IN PAS-7).
N-LINKED (GLCNAC. . .) (HYBRID) (IN PAS-6
AND PAS-7).
N-LINKED (GLCNAC. . .) (HIGH MANNOSE) (IN
PAS-6).
MISSING (IN SHORT ISOPFORM).
A -> F (IN REF. 1).
L -> Q (IN REF. 1).
Query Match 43.4%; Score 1159.5; DB 1; Length 427;
Best local similarity 46.7%; Pred. No. 8.8e-81;
Matches 214; Conservative 74; Mismatches 115; Indels 55; Gaps 5;
QY 23 GDICNPENCNGICLSLAD---SFCPCPGFPGMPCSSVEYASDEEKPFSAGPCI 79
DB 21 GDCDSLCILHGGTCL-LNEDRTPPYCLCPBGFGLCNE-----TEHGCF 67
QY 80 PNCNGGTCEIS-EAYRGDTFYGVCPRGNGICGHNINEEAEPFCRNGICITDLY 138
DB 68 PNCNDHDECCQVTDSDSHRGDFIYICKCPGLGYGHCEFT----- 108
QY 139 ANYSCPCPEFGMRCQKCSGHGIGGIIISNOQTIVASNNHAFGLQKPPYARLNK 198
DB 109 -----CTSPICMOTGAIADSQISASMSMLGEGIQARAPELARIHQ 149
QY 199 KGLINAKTRAENDRPWQIMLQKRAVYTGQAKRIGSPETIKSYKIATVANDGKTWA 258
DB 150 TGIYNAWTSIGNDKNPWLOVNLKRMVTVGVQASRGASAEYLTKEVAVSTDGROFO 209
QY 259 MYKVKGTNEEMVFRGNVDNNTPYANSETPPIKAQYVRLYPOICRRHCTLMELLCCEL 318
DB 210 FIVAGRSQDKIFFIGNVNSGLKINLEDFLETOYVRLVPIICRCTLREFFLLCCEING 269
QY 319 GSPVPGMSKGIHODVQIVTASSVFRTLMMDFTWEPKRRARDGQKVNMTSGHNDOSWL 378
DB 270 CTEPLGKNTITPNQITASSYVTKWLSAFSNFPYARLDNCGKNMTAGTNSASML 329
QY 379 QVDLPLVKYTGITIGKAGFGHVOYGVSTKVLASNDGEHMYHODEKORAKYFQGNFD 438
DB 330 QIDLSQKRVGTGITIGARGHGIQVAAVRAVAGDGVLTWEYKDPGASSEKIFPGMD 389
QY 439 NDTIRKNVDPPIYARFIRILPMASMYGRTILRELLGCG 476
DB 390 NNSHKNIETEPFOARFVRIDPVAMHNRITLRELLGCG 427
RESULT 5
MFGM_HUMAN STANDARD; PRT; 387 AA.
ID MFGM_HUMAN
AC Q08431;
DT 01-OCT-1996 (Rel. 34, Created)

DT 01-NOV-1997 (rel. 35, last sequence update)
 DT 30-MAY-2000 (rel. 39, last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MEG-E8) (HMEG)
 DE (BREAST EPITHELIAL ANTIGEN BA46) (MEGM) [CONTAINS: MEDIN].
 GN MEG8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Breast and Breast carcinoma;
 RX MEDLINE=96213908; PubMed=8639264;
 RA Couto J.R., Taylor M.R., Godwin S.G., Ceriani R.L., Peterson J.A.;
 RT "Cloning and sequence analysis of human breast epithelial antigen
 BA46 reveals an RGD cell adhesion sequence presented on an epidermal
 growth factor-like domain.";
 RT DNA Cell Biol. 15:281-286(1996).
 RL [2]
 RP SEQUENCE OF 170-387 FROM N.A.
 RC TISSUE-Mammary gland;
 RX MEDLINE=91371351; PubMed=1909932;
 RA Larocca D., Peterson J.A., Ureia R., Kuniyoshi J., Bistrain A.M.,
 RA Ceriani R.L.;
 RT "A Mr 46,000 human milk fat globule protein that is highly expressed
 in human breast tumors contains factor VIII-like domains.";
 RT Cancer Res. 51:4994-4998(1991).
 RL [3]
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RC TISSUE-Milk;
 RX GLUFFIDE M.G., Cavaleto M., Giunta C., Conti A.,
 RA Godovac-Zimmermann J.;
 RT "Isolation and characterization of full and truncated forms of human
 breast carcinoma protein BA46 from human milk fat globule membranes.";
 RT J. Protein Chem. 17:143-148(1998).
 RL [4]
 RP SEQUENCE OF 268-317, AND IDENTIFICATION OF MEDIN.
 RX MEDLINE=99342076; PubMed=10411933;
 RA Haeggevisat B., Naeslund J., Sletten K., Westermarck G.T., Mucchiano G.,
 RA Thernberg L.O., Nordstedt C., Engstrom U., Westermarck P.;
 RT "Medin: an integral fragment of aortic smooth muscle cell-produced
 lactadherin forms the most common human amyloid.";
 RT Proc. Natl. Acad. Sci. U.S.A. 96:8669-8674(1999).
 RL [5]
 RP CHARACTERIZATION.
 RX MEDLINE=97405885; PubMed=9260929;
 RA Taylor M.R., Couto J.R., Scallan C.D., Ceriani R.L., Peterson J.A.;
 RT "Lactadherin (formerly BA46), a membrane-associated glycoprotein
 expressed in human milk and breast carcinomas, promotes Arg-Gly-Asp
 (RGD)-dependent cell adhesion.";
 RT DNA Cell Biol. 16:661-669(1997).
 RL [6]
 CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. BINDS
 SPECIFICALLY TO ROTAVIRUS AND INHIBITS ITS REPLICATION.
 CC -1- FUNCTION: MEDIN IS THE MAIN CONSTITUENT OF AORTIC MEDIAL AMYLOID.
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND AORTIC
 MEDIA. OVEREXPRESSED IN SEVERAL CARCINOMAS.
 CC -1- PTM: MEDIN HAS A RAGGED N-TERMINUS WITH MINOR SPECIES STARTING AT
 AMINO ACID 264 AND 273.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, U58516; AAC50549.1; -
 DR EMBL, S56151; AAB19771.1; -

DR MW: 602281; -
 DR InterPro: IPR000421; -
 DR InterPro: IPR000561; -
 DR Pfam: PF00754; EGF_1.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01285; FAS5C_1; 2.
 DR PROSITE: PS01286; FAS5C_2; 2.
 DR Signal: Glycoprotein; Milk; Repeat; EGF-like domain; Amyloid.
 KW SIGNAL
 FT CHAIN 1 23
 FT CHAIN 24 387
 FT CHAIN 202 387
 FT CHAIN 268 317
 FT DOMAIN 24 67
 FT DOMAIN 70 225
 FT DOMAIN 230 387
 FT SITE 46 48
 FT DISULFID 27 38
 FT DISULFID 32 55
 FT DISULFID 57 66
 FT DISULFID 70 225
 FT DISULFID 212 216
 FT DISULFID 230 387
 FT CARBOHYD 238 238
 FT CARBOHYD 325 325
 FT CARBOHYD 329 329
 FT CARBOHYD 350 350
 SQ SEQUENCE 387 AA; 43123 MW; 2EB571DEC83782D CnC64;
 LACTADHERIN, SHORT FORM.
 MEDIN.
 EGF-LIKE.
 F5/8 TYPE C 1.
 F5/8 TYPE C 2.
 CELL ATTACHMENT SITE (POTENTIAL).
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 38.3%; Score 1023; DB 1; Length 387;
 Best local Similarity 48.9%; Pred. No. 1.7e-70;
 Matches 196; Conservative 55; Mismatches 108; Indels 42; Gaps 4;
 QY 78 CIPNPGHNGTC-EISEATRGDTFGICVCKPRGFNGHICQININCEAPECRNGGICTD 136
 DB 27 CSKNPCHNGGLEIEISQEVKGVFPSTCTCLKAGNN----- 65
 QY 137 LVANYSCECGEPGMRNCKYKSGHLIGGILISNOQITASSNHRALFGLQWYPYARL 196
 DB 66 -----CEKCEPELGMNGNANSQIASSVRVFLGLQHWVELARL 108
 QY 197 NKGGLINAWTAENDRWPIQIMLQKKRMYGVYTGQAKRISPEYIKYKIAISDQKT 256
 DB 109 NRGAVNAVMPSSNDNPQIVMLIRRMVTVGVVTOGASRLASHLEYLAKFVAYSLNGE 168
 QY 257 W-AMKYKKGTEEMVNRGVNDNTPYANSFTPRKQVRYLYPOICRRCTLMELIGE 315
 DB 169 FDFTHDNKKKKEFV--GNNKNAVHVNLFEPVEQVRYLYPSCHTACTLRFELIGE 226
 QY 316 LSGCEPLGKSGCHIDYOITASSVFTLMDFTWEPKRAILDKQGVNMTSGHNDOS 375
 DB 227 LNCANPLGLKNNISIPDKQITASSYKWTGLHLEFMSNPYARLDKGNFNANWAGSYGND 286
 QY 376 QWLOVDLVLPFKYTGITGQAKPGHVPQVSGYKRLAYSNDGEHMMHODEKQKDKVPOG 435
 DB 287 QWLOVDLVGSSKEVYTGITGQARNFGSVQFVASYKVAISNSAWNTVEYDPRGSSKIFPG 346
 QY 436 NQMDTHRRKVVDPPIYARFIRILPMSWYGRITLRSLELGC 476
 DB 347 NMDNHSKKNLFEPIIARVIRILPYAMHRIALRELELGC 387
 RESULT 6
 ID PA5_HUMAN STANDARD; PRT; 2224 AA.
 AC P12259; Q14285;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-FEB-1994 (Rel. 28, last sequence update)
 DT 01-OCT-2000 (Rel. 40, last annotation update)
 DE COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).
 GN F5.

OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92232668; PubMed=1567832;
 RA Cripe L.D., Moore K.D., Kane W.H.;
 RT "Structure of the gene for human coagulation factor V.";
 RL Biochemistry 31:3777-3785(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87260886; PubMed=3110773;
 RA Jenny R.J., Pittman D.D., Tootle J.J., Kriz R.W., Aldape R.A.,
 RA Hewick R.M., Kaufman R.J., Mann K.G.;
 RT "Complete cDNA and derived amino acid sequence of human factor V.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987).
 RN [3]
 RP SEQUENCE OF 1-1600 FROM N.A.
 RA MEDLINE=88107560; PubMed=2827731;
 RA Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;
 RT "Cloning of cDNAs coding for the heavy chain region and connecting
 RT region of human factor V, a blood coagulation factor with four types
 RT of internal repeats.";
 RL Biochemistry 26:6508-6514(1987).
 RN [4]
 RP SEQUENCE OF 1168-1215 AND 1315-2224 FROM N.A.
 RA MEDLINE=86313665; PubMed=3092220;
 RA Kane W.H., Davie E.W.;
 RT "Cloning of a cDNA coding for human factor V, a blood coagulation
 RT factor homologous to factor VIII and ceruloplasmin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).
 RN [5]
 RP PARTIAL SEQUENCE FROM N.A.
 RA TISSUE=Fibroblast;
 RA MEDLINE=93203619; PubMed=8454869;
 RA Shen N.L.L., Fan S.-T., Pyati J., Graff R., Lapolla R.J.,
 RA Edgington T.S.;
 RT "The serine protease cofactor factor V is synthesized by
 RT lymphocytes.";
 RL J. Immunol. 150:2992-3001(1993).
 RN [6]
 RP VARIANT APCR GIN-534.
 RA MEDLINE=94217810; PubMed=8164741;
 RA Bertina R.M., Koelman B.P.C., Koster T., Rosendaal F.R.,
 RA Dirven R.J., de Ronde H., van der Velde P.A., Reitsma P.H.;
 RT "Mutation in blood coagulation factor V associated with resistance to
 RT activated protein C.";
 RL Nature 369:64-67(1994).
 CC -1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES
 CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.
 CC -1- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT
 CC CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS
 CC IS CALCIUM-DEPENDENT.
 CC -1- DOMAIN: DOMAIN B CONTAINS 35 X 9 AA TANDEM REPEATS, AND 2 X 17 AA
 CC REPEATS.
 CC -1- PTM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE
 CC COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-
 CC TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).
 CC -1- DISEASE: OMEN PARAHEMOPHILIA, AN HEMORRHAGIC DIATHESIS, IS DUE
 CC TO A DEFICIENCY OF FACTOR V. OTHER DEFECTS IN F5 RESULTS IN A
 CC FORM OF THROMBOPHILIA KNOWN AS APC RESISTANCE (APCR). THE APCR
 CC MUTATION IS FOUND IN ABOUT 5% OF THE POPULATION WHICH SUGGEST THAT
 CC A SLIGHT THROMBOTIC TENDENCY MAY CONFER SOME ADVANTAGE IN FETAL
 CC IMPLANTATION.
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS, EACH IS COMPOSED OF
 CC 2 PLASTOCYANIN-LIKE REPEATS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 CC DR EMBL, L32779; AAB59401.1; -
 CC DR EMBL, L32755; AAB59401.1; JOINED.
 CC DR EMBL, L32756; AAB59401.1; JOINED.
 CC DR EMBL, L32757; AAB59401.1; JOINED.
 CC DR EMBL, L32758; AAB59401.1; JOINED.
 CC DR EMBL, L32759; AAB59401.1; JOINED.
 CC DR EMBL, L32760; AAB59401.1; JOINED.
 CC DR EMBL, L32761; AAB59401.1; JOINED.
 CC DR EMBL, L32762; AAB59401.1; JOINED.
 CC DR EMBL, L32763; AAB59401.1; JOINED.
 CC DR EMBL, L32764; AAB59401.1; JOINED.
 CC DR EMBL, L32765; AAB59401.1; JOINED.
 CC DR EMBL, L32766; AAB59401.1; JOINED.
 CC DR EMBL, L32767; AAB59401.1; JOINED.
 CC DR EMBL, L32768; AAB59401.1; JOINED.
 CC DR EMBL, L32769; AAB59401.1; JOINED.
 CC DR EMBL, L32770; AAB59401.1; JOINED.
 CC DR EMBL, L32771; AAB59401.1; JOINED.
 CC DR EMBL, L32772; AAB59401.1; JOINED.
 CC DR EMBL, L32773; AAB59401.1; JOINED.
 CC DR EMBL, L32774; AAB59401.1; JOINED.
 CC DR EMBL, L32775; AAB59401.1; JOINED.
 CC DR EMBL, L32776; AAB59401.1; JOINED.
 CC DR EMBL, L32777; AAB59401.1; JOINED.
 CC DR EMBL, L32778; AAB59401.1; JOINED.
 CC DR EMBL, L16967; AAB52424.1; -
 CC DR EMBL, M14335; AAB59532.1; -
 CC DR PIR, A25897; A25897.
 CC DR PIR, A28028; A28028.
 CC DR HSSP, P00A50; 1KCM.
 CC DR MIM, L34400; -
 CC DR MIM, L88055; -
 CC DR MIM, L227310; -
 CC DR MIM, L227400; -
 CC DR InterPro: IPR000421; -
 CC DR InterPro: IPR001117; -
 CC DR Pfam: PF00394; Cu-oxidase; 3.
 CC DR Pfam: PF00754; F5_F8_type_C; 2.
 CC DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 2.
 CC DR PROSITE: PS01285; FA58C_1; 2.
 CC DR PROSITE: PS01286; FA58C_2; 2.
 CC KW Blood coagulation; Plasma; Glycoprotein; Calcium; Signal; Zymogen;
 CC Repeat; Polymorphism; Disease mutation; Thrombophilia.
 CC FT SIGNAL 1 28
 CC FT CHAIN 29 2224
 CC FT CHAIN 29 737
 CC FT PEPTIDE 738 1573
 CC FT CHAIN 1574 2224
 CC FT DOMAIN 30 329
 CC FT DOMAIN 30 193
 CC FT DOMAIN 203 329
 CC FT DOMAIN 348 684
 CC FT DOMAIN 348 526
 CC FT DOMAIN 536 684
 CC FT DOMAIN 692 1573
 CC FT DOMAIN 895 928
 CC FT REPEAT 895 911
 CC FT REPEAT 912 928
 CC FT REPEAT 913 1148
 CC FT SIMILAR 1185 1463
 CC FT DOMAIN 1185 1463
 CC FT REPEAT 1185 1193
 CC FT REPEAT 1194 1202
 CC FT REPEAT 1203 1211
 CC FT REPEAT 1212 1220
 CC FT REPEAT 1221 1229
 CC FT REPEAT 1230 1238
 CC FT REPEAT 1239 1247
 CC FT REPEAT 1247 1247
 CC -----
 CC 1. COAGULATION FACTOR V.
 CC 2. HEAVY CHAIN.
 CC 3. ACTIVATION PEPTIDE (CONNECTING REGION).
 CC 4. LIGHT CHAIN.
 CC 5. F5/8 TYPE A 1.
 CC 6. PLASTOCYANIN-LIKE 1.
 CC 7. PLASTOCYANIN-LIKE 2.
 CC 8. F5/8 TYPE A 2.
 CC 9. PLASTOCYANIN-LIKE 3.
 CC 10. PLASTOCYANIN-LIKE 4.
 CC 11. B.
 CC 12. 2 X 17 AA TANDEM REPEATS.
 CC 13. 1.
 CC 14. TO 14 AA REPEATS IN BOVINE FAS.
 CC 15. 35 X 9 AA TANDEM REPEATS OF [TNP]-L-S-P-
 CC D-L-S-Q-T (APPROXIMATE).
 CC 16. 1.
 CC 17. 2.
 CC 18. 3.
 CC 19. 4.
 CC 20. 5.
 CC 21. 6.
 CC 22. 7.

```

FT REPEAT 1248 1256 8.
FT REPEAT 1257 1265 9.
FT REPEAT 1266 1274 10.
FT REPEAT 1275 1283 11.
FT REPEAT 1284 1292 12.
FT REPEAT 1293 1301 13.
FT REPEAT 1302 1310 14.
FT REPEAT 1311 1319 15.
FT REPEAT 1320 1328 16.
FT REPEAT 1329 1337 17.
FT REPEAT 1338 1346 18.
FT REPEAT 1347 1355 19.
FT REPEAT 1356 1364 20.
FT REPEAT 1365 1373 21.
FT REPEAT 1374 1382 22.
FT REPEAT 1383 1391 23.
FT REPEAT 1392 1400 24.
FT REPEAT 1401 1409 25.
FT REPEAT 1410 1418 26.
FT REPEAT 1419 1427 27.
FT REPEAT 1428 1436 28.
FT REPEAT 1437 1445 29.
FT REPEAT 1446 1454 30.
FT REPEAT 1455 1463 31.
FT REPEAT 1464 1472 32.
FT REPEAT 1473 1481 33.
FT REPEAT 1482 1490 34.
FT REPEAT 1493 1501 35.
FT DOMAIN 1578 1586 36.
FT DOMAIN 1578 1586 37.
FT DOMAIN 1578 1586 38.
FT DOMAIN 1578 1586 39.
FT DOMAIN 1578 1586 40.
FT DOMAIN 1578 1586 41.
FT DOMAIN 1578 1586 42.
FT DOMAIN 1578 1586 43.
FT DOMAIN 1578 1586 44.
FT DOMAIN 1578 1586 45.
FT DOMAIN 1578 1586 46.
FT DOMAIN 1578 1586 47.
FT DOMAIN 1578 1586 48.
FT DOMAIN 1578 1586 49.
FT DOMAIN 1578 1586 50.
FT DOMAIN 1578 1586 51.
FT DOMAIN 1578 1586 52.
FT DOMAIN 1578 1586 53.
FT DOMAIN 1578 1586 54.
FT DOMAIN 1578 1586 55.
FT DOMAIN 1578 1586 56.
FT DOMAIN 1578 1586 57.
FT DOMAIN 1578 1586 58.
FT DOMAIN 1578 1586 59.
FT DOMAIN 1578 1586 60.
FT DOMAIN 1578 1586 61.
FT DOMAIN 1578 1586 62.
FT DOMAIN 1578 1586 63.
FT DOMAIN 1578 1586 64.
FT DOMAIN 1578 1586 65.
FT DOMAIN 1578 1586 66.
FT DOMAIN 1578 1586 67.
FT DOMAIN 1578 1586 68.
FT DOMAIN 1578 1586 69.
FT DOMAIN 1578 1586 70.
FT DOMAIN 1578 1586 71.
FT DOMAIN 1578 1586 72.
FT DOMAIN 1578 1586 73.
FT DOMAIN 1578 1586 74.
FT DOMAIN 1578 1586 75.
FT DOMAIN 1578 1586 76.
FT DOMAIN 1578 1586 77.
FT DOMAIN 1578 1586 78.
FT DOMAIN 1578 1586 79.
FT DOMAIN 1578 1586 80.
FT DOMAIN 1578 1586 81.
FT DOMAIN 1578 1586 82.
FT DOMAIN 1578 1586 83.
FT DOMAIN 1578 1586 84.
FT DOMAIN 1578 1586 85.
FT DOMAIN 1578 1586 86.
FT DOMAIN 1578 1586 87.
FT DOMAIN 1578 1586 88.
FT DOMAIN 1578 1586 89.
FT DOMAIN 1578 1586 90.
FT DOMAIN 1578 1586 91.
FT DOMAIN 1578 1586 92.
FT DOMAIN 1578 1586 93.
FT DOMAIN 1578 1586 94.
FT DOMAIN 1578 1586 95.
FT DOMAIN 1578 1586 96.
FT DOMAIN 1578 1586 97.
FT DOMAIN 1578 1586 98.
FT DOMAIN 1578 1586 99.
FT DOMAIN 1578 1586 100.

```

Query Match 25.8%, Score 689, DB 1, Length 2224;
 Best Local Similarity 42.7%, Pred. No. 2, de-44;
 Matches 144; Conservative 55; Mismatches 122; Indels 16; Gaps 4;

```

QY 144 ECGEFGNRCQKCSGHLGEGIGIISNOQTASNNHRLFGLOKVPYARLKKGLIN 203
DB 1897 QTFPLMDNDCHMP-----MGLSTGIISDSQIKASE-----FLGWPRRLARLNNGSYN 1946
QY 204 AMT-----AAENBRPMIQTINLQKMRVTGVTTCAGAKRIGSEPTKSKIAVSDGKTWAM 259
DB 1947 AWSGVERLAEPASKEPIQVDMQKEVITIGTQOGAKHLYKSCYTFTEFYVAYSSQINWQI 2006
QY 260 YKKGKTNEEVRFGVNDNTPRANSFPPIKAOVYRLPOLCRHCHLRLMELGCELSGC 319
DB 2007 FKGNSTRNMYRNGNSDASTIKENOPDPPIAARIYRISPTAYAKRPLRLLELOGCEVNGC 2066
QY 320 SEPLGKSGHIDYOITASSVFTLNDMTPEPRKARLDKOGKVNMTSGHNDQSOWLO 379
DB 2067 STEPLGNGKIEKQITASSFFKSWMGDY--WEPRFALNAGVHNMOKAKANNKQWLE 2124
QY 380 VDLVLPKATGIIITOGAFGHVOPVGSYKLAYSNDCEHHWVHODEQORDKVEGQEDN 439
DB 2125 IDLLKIKKITAIIITOGCKLSSEMTYKSTYIHSQGVEMKPYLKSMDKIFEGTNT 2184
QY 440 DTHRRKNVIDPIYARFIRLLPMSWYGRITLSELGCG 476

```

```

DB 2185 KGHVKNFNPPIISRFIRVIRPKTMQSTILRLLELGC 2221
RESULT 7
PAB_MOUSE STANDARD; PRT: 2319 AA.
ID PAB_MOUSE
AC Q06194;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT).
GN CFB OR F8C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=93300511; Pubmed=8314577;
RA Elder B., Lachic D., Glitscher J.;
RT "Sequence of the murine factor VIII cDNA."
RL Genomics 16:374-379(1993).
CC -1- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
CC ACTIVATED FORM, FACTOR XA.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
CC 2 PLASTOCYANIN-LIKE REPEATS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: L05573; AAA37385.1; -.
CC PIR: A47004; A47004.
CC HSP: P00451; 1CFG.
CC MGD: MGI:88383; F8.
CC InterPro: IPR000421; -.
CC InterPro: IPR001117; -.
CC Pfam: PF00394; Cu-oxidase; 3.
CC Pfam: PF00754; F5_F8_type_C; 2.
CC PROSITE: PS00079; MULTICOOPER_OXIDASE1; 3.
CC PROSITE: PS01285; FA58C_1; 2.
CC PROSITE: PS01286; FA58C_2; 2.
CC Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
CC Signal; Glycoprotein; Sulfatation.
CC SIGNAL 1 19
CC CHAIN 20 2319
CC DOMAIN 20 349 F5/8 TYPE A 1.
CC DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
CC DOMAIN 207 349 PLASTOCYANIN-LIKE 2.
CC DOMAIN 399 730 F5/8 TYPE A 2.
CC DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
CC DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
CC DOMAIN 760 1640 B.
CC DOMAIN 1683 2008 F5/8 TYPE A 3.
CC DOMAIN 1683 1845 PLASTOCYANIN-LIKE 5.
CC DOMAIN 1855 2008 PLASTOCYANIN-LIKE 6.
CC DOMAIN 2008 2156 F5/8 TYPE C 1.
CC DOMAIN 2161 2313 F5/8 TYPE C 2.
CC SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
CC SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
CC SITE 1678 1679 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
CC SITE 1324 1325 CLEAVAGE (ACTIVATION) (BY SIMILARITY).

```

```

FT SITE 1640 1641 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
FT MOD_RES 367 367 SUFFATATION (BY SIMILARITY).
FT MOD_RES 737 737 SUFFATATION (BY SIMILARITY).
FT MOD_RES 738 738 SUFFATATION (BY SIMILARITY).
FT MOD_RES 742 742 SUFFATATION (BY SIMILARITY).
FT MOD_RES 1669 1669 SUFFATATION (REQUIRED FOR VWF BINDING)
FT MOD_RES 1669 1669 (BY SIMILARITY)
FT MOD_RES 1687 1687 SUFFATATION (BY SIMILARITY).
FT DISULFID 173 199 PROBABLE.
FT DISULFID 347 373 PROBABLE.
FT DISULFID 1819 1845 PROBABLE.
FT DISULFID 2008 2156 BY SIMILARITY.
FT DISULFID 2161 2313 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1136 1136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1161 1161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1192 1192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1255 1255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1274 1274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1302 1302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1316 1316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1340 1340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1797 1797 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2105 2105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2319 2319 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 2319 AA; 266148 MW; FD054DE051DB2A01 CRC64;

Query Match 25.3%; Score 675; DB 1; Length 2319;
Best Local Similarity 44.6%; Pred. No. 2,9e-43;
Matches 154; Conservative 42; Mismatches 119; Indels 30; Gaps 6;

```

```

ID FAS_BOVIN STANDARD; PRT; 2211 AA.
AC Q28107; Q28108;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).
GN F5.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=liver;
RX MEDLINE=92147638; PubMed=1737753;
RA Guitto E.R., Ramon C.T., Mann K.G., Maggillivray R.T.;
RT "The complete cDNA sequence of bovine coagulation factor V.";
RL J. Biol. Chem. 267:29712-29718(1992).
CC -1- FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES
CC -1- WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.
CC -1- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT
CC -1- CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS
CC -1- IS CALCIUM-DEPENDENT.
CC -1- DOMAIN: DOMAIN B CONTAINS 29.5 X 9 AA TANDEM REPEATS, AND 2 X 17
CC -1- AA REPEATS.
CC -1- PTM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE
CC -1- COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-
CC -1- TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).
CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
CC -1- 2 PLASTOCYANIN-LIKE REPEATS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81440; AAA30512.1; -
DR EMBL; M81441; AAA30513.1; -
DR HSSP; P00450; 1KCM.
DR InterPro; IPR000421; -
DR InterPro; IPR001117; -
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5_F8_typeC; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
KW Blood coagulation; Plasma; Glycoprotein; Calcium; Signal; Zymogen;
KW Repeat.
FT SIGNAL. 1 28
FT CHAIN. 29 2211
FT CHAIN. 29 741
FT PEPTIDE 742 1564
FT CHAIN 1565 2211
FT DOMAIN 30 327
FT DOMAIN 30 193
FT DOMAIN 203 327
FT DOMAIN 348 686
FT DOMAIN 348 525
FT DOMAIN 535 686
FT DOMAIN 535 1564
FT DOMAIN 696 915
FT SIMILAR 899 915
FT DOMAIN 1124 1151
FT DOMAIN 1124 1137
FT REPEAT 1138 1151
FT REPEAT 1188 1453
FT DOMAIN 1188 1453

```

RESULT 8
FAS_BOVIN


```

FT REPEAT 1188 1196 1.
FT REPEAT 1197 1205 2.
FT REPEAT 1206 1214 3.
FT REPEAT 1215 1223 4.
FT REPEAT 1224 1232 5.
FT REPEAT 1233 1241 6.
FT REPEAT 1242 1250 7.
FT REPEAT 1251 1259 8.
FT REPEAT 1260 1268 9.
FT REPEAT 1269 1277 10.
FT REPEAT 1278 1286 11.
FT REPEAT 1287 1295 12.
FT REPEAT 1296 1304 13.
FT REPEAT 1305 1313 14.
FT REPEAT 1314 1322 15.
FT REPEAT 1323 1331 16.
FT REPEAT 1332 1340 17.
FT REPEAT 1341 1349 18.
FT REPEAT 1350 1358 19.
FT REPEAT 1359 1367 20.
FT REPEAT 1368 1376 21.
FT REPEAT 1377 1385 22.
FT REPEAT 1386 1394 23.
FT REPEAT 1395 1403 24.
FT REPEAT 1404 1412 25.
FT REPEAT 1413 1421 26.
FT REPEAT 1422 1430 27.
FT REPEAT 1431 1439 28.
FT REPEAT 1440 1448 29. (PARTIAL).
FT REPEAT 1445 1453 30.
FT DOMAIN 1569 1890 F5/8 TYPE A 3.
FT DOMAIN 1748 1890 PLASTOCYANIN-LIKE 5.
FT DOMAIN 1894 2048 F5/8 TYPE C 1.
FT DOMAIN 2053 2208 F5/8 TYPE C 2.
FT SITE 741 742 CLEAVAGE (BY THROMBIN)
FT SITE 1034 1035 CLEAVAGE (BY THROMBIN)
FT SITE 1564 1565 CLEAVAGE (BY THROMBIN)
FT DISULFID 167 193 PROBABLE.
FT DISULFID 499 525 PROBABLE.
FT DISULFID 1712 1738 PROBABLE.
FT DISULFID 1894 2048 BY SIMILARITY.
FT DISULFID 2053 2208 BY SIMILARITY.
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 756 756 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 902 902 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1053 1053 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1062 1062 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1094 1094 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1451 1451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1490 1490 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1550 1550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1690 1690 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1839 1839 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1997 1997 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2196 2196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 587 592 NPTIPA -> T (IN VARIANT 2).
SQ SEQUENCE 2211 AA: 246981 MW: CBBP90B736667C45 CRC64;

```

```

Query Match 25.1%; Score 671; DB 1; Length 2211;
Best Local Similarity 42.9%; Pred. No. 5.6e-43;
Matches 139; Conservative 55; Mismatches 118; Indels 12; Gaps 5;

QY 157 KCSGHGIGEGIIISNOOTASSNHAFGLQKWPYARLNKKGLINAWA----AENDR 212
DB 1893 ECKMPWGLSTGLADSOADSE---FWGY--WEPKARLNGSGSYNAWAEKLSFEFNP 1946

QY 213 WPMIQTINLRKMRVYGVITQGAKRIGSPYIKSYKIAYSDNGKTMAMKYKGTNEEMVR 272
DB 1947 EPWIDQMOKEVLTGIGTQGAHLYKPYTTEFCVAYSIDRKMRWLFKGNSTRANMYEG 2006

QY 273 GNVDNNTPYANSPTPIRPAOVRYLPOICRRHCTRLMELGCELSGCSPELGKSGHIDD 332
DB 2007 GNSDASTIKENQIDPPVARIKISPTGSYKPKPALRLDQCEVNGCSTPLGMEGSKLEN 2066

QY 333 YQTTASSVERTLNMDFTEWPRKARLDKQKYNAMTSGHNDQSWLQVLDLVPTKVTGII 392
DB 2067 KOITASS-FKKSWMGNY-WEPLARLNAQGHVNAQKANNNNQWLOIDLTIKKITTAVI 2124

QY 393 TOGAKDFGHVQFVSGYKLAYSNDGEHMMVHODEKQKDKYQGFNDPRKKNVIDPPIY 452
DB 2125 TQCKSLSEMYKSTTIHSDQGTDMKPYREKSSMTDKIFEKGNNNYRGVKNFNPPII 2184

QY 453 ARPIRLPWSYGRITLRSLELGC 476
DB 2185 SRFRIIPKTPWQSIALRLLEFGC 2208

RESULT 9
FAB_HUMAN STANDARD; PRT: 2351 AA.
AC P00451;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT)
DE (ANTIHEMOPHILIC FACTOR) (AHF).
GN F8 OR F8C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86081164; PubMed=3935400;
RA Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA Hartog K., Kuo C.H., Mastarz F.R., Merryweather J.P., Najarian R.,
RA Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA Tudea M.S., Valenzuela P., Dahl H.-H.M., Favalato J., Hansen J.,
RA Nordfang O., Ezban M.;
RT Characterization of the polypeptide composition of human factor
RT VIII:C and the nucleotide sequence and expression of the human kidney
RT cDNA.
RL DNA 4:333-349(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061548; PubMed=6438526;
RA Wood W.T., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA Keyt B., Seeburg P.H., Smith D.H., Hollinshead F., Wion K.L.,
RA Delwart E., Tuddenham E.G.D., Vohar G.A., Lawn R.M.;
RT Expression of active human factor VIII from recombinant DNA clones.
RL Nature 312:330-337(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061550; PubMed=6438528;
RA Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,
RA Pittman D.D., Kauman R.J., Brown E., Shoemaker C., Orr E.C.,
RA Applebitt G.W., Foster W.B., Coe M.L., Knutson G.J., Pass D.N.,
RA Hewick R.M.;
RT Molecular cloning of a cDNA encoding human antihemophilic factor.
RL Nature 312:342-347(1984).

```

RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93265012; PubMed-1303178;
 RA Gitschier J., Wood W.I.;
 RT "Sequence of the exon-containing regions of the human factor VIII
 RT gene";
 RL Hum. Mol. Genet. 1:199-200(1992).
 RN [5]
 RP SEQUENCE OF 2064-2070 FROM N.A.
 RA de Water N.S., Williams R., Browett P.J.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SULFATATION OF 1699.
 RX MEDLINE-91093266; PubMed-1898735;
 RA Leyte A., van Schilpeld H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
 RA Merens K., van Mourik J.A.;
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RP SULFATATION.
 RX MEDLINE-92207952; PubMed-1554716;
 RA Pittman D.B., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX MEDLINE-95200924; PubMed-7893714;
 RA Gilbert G.E., Baleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE-91221499; PubMed-1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE-89088506; PubMed-2491949;
 RA White G.C.II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE-95245332; PubMed-7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT GLN-2326.
 RX MEDLINE-86233434; PubMed-3012775;
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophilic";
 RL Science 233:1415-1416(1986).
 RN [13]
 RP VARIANT PRO-2135.
 RX MEDLINE-88096539; PubMed-3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT GLN-2228.
 RX MEDLINE-88191889; PubMed-2833855;
 RA Yousoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides";

RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT GLY-291.
 RX MEDLINE-88220354; PubMed-2835904;
 RA Yousoufian H., Wong C., Aronis S., Platakoukis H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu->Gly substitution
 RT in exon 7 of the factor VIII gene";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT CYS-1708.
 RX MEDLINE-89274393; PubMed-2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A";
 RL Blood 73:2117-2122(1989).
 RN [17]
 RP VARIANT CYS-391.
 RX MEDLINE-90001543; PubMed-2506948;
 RA Shima M., Ware J., Yoshioke A., Fukui H., Pulcher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule";
 RL Blood 74:1612-1617(1989).
 RN [18]
 RP VARIANT LEU-189.
 RX MEDLINE-90057680; PubMed-2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII:C gene
 RT resulting in moderately severe hemophilia A";
 RL Blood 74:2688-2691(1989).
 RN [19]
 RP VARIANT LEU-2326.
 RX MEDLINE-89197216; PubMed-2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg->Leu substitution in exon 26
 RT of the factor VIII gene";
 RL Hum. Genet. 81:335-338(1989).
 RN [20]
 RP VARIANT HIS-391.
 RX MEDLINE-89264602; PubMed-2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372->histidine)";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [21]
 RP VARIANT CYS-1708.
 RX MEDLINE-90105723; PubMed-2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A";
 RL Blood 75:384-389(1990).
 RN [22]
 RP VARIANT GLN-2228 AND LEU-2326.
 RX MEDLINE-90123183; PubMed-2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent";
 RL Blood 75:662-670(1990).
 RN [23]
 RP VARIANT CYS-391.
 RX MEDLINE-90329422; PubMed-1973901;
 RA Patinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ haemophilia A due to a missense mutation (372->Cys) at the
 RT internal heavy chain thrombin cleavage site";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [24]
 RP VARIANTS PHE-1699 AND CYS-1708.

FT CONFLICT 792 792 G -> Q (IN REF. 2).
 FT CONFLICT 1133 1133 E -> F (IN REF. 2).
 FT CONFLICT 1191 1191 I -> L (IN REF. 2).
 FT CONFLICT 1209 1209 R -> F (IN REF. 2).
 FT CONFLICT 1437 1437 C -> G (IN REF. 2).
 FT CONFLICT 1456 1456 F -> R (IN REF. 2).
 FT CONFLICT 1539 1539 F -> R (IN REF. 2).
 FT CONFLICT 1546 1546 Q -> N (IN REF. 2).
 SQ SEQUENCE 2133 AA; 239304 MM; 152BBA8997F570DA CRC64;

Query Match 23 8% Score 636.5; DB 1; Length 2133;
 Best Local Similarity 40.6%; Pred No. 2,2e-40;
 Matches 146; Conservative 49; Mismatches 136; Indels 29; Gaps 6;

DB 122 EC-EAEPGRNGICTDLVANYSCCEPGEFGRNCGKSGHLGIGGIIISNOQTASSNH 180
 DB 1800 ECLIGELHQAQMSTFTFLVYSKECOAP-----LGMASGRIRFOQTASSGOY 1844
 QY 181 RALEGLQWYFYARLNKGLINAMTAENDMPWIOINLORKMRYGVITOGAKRIGSP 240
 DB 1845 -----GQWAPKLARLHYSISINAMST--KDPHSMTIKVDLLAPMIIHGIMTQGAOKFSS 1896
 QY 241 EYISYKTAISNDGKTAAKVKYGTNEEYVFGNDNNTPPANSTPPIRAKOYRLPQI 300
 DB 1897 LYISOFITIMSYLDRNMQSTRGSTGLMVFEGNDASGKHNTFNPPYARVIRLHPTR 1956
 QY 301 CRRHCTLMELLAGCELSGCEPLGKMSGHIQDYQITASSVFTLNMDFWTEPRKARLDK 360
 DB 1957 YSINSTLMEIIMGCDLNSCMPLGOMKKAISDSQITASS--HLSNIFATWSPSOAHL 2013
 QY 361 OGKNAVMTSGHNDQSOHLOYDLLVPTKVTGILTOGAKPGHVOFGSKLAYSNDGHHM 420
 DB 2014 OGRINAMPRVSSAEMLQYDLQKTYKVTGITTQGVKSLISSMYKEFLVSSSDGGRWT 2073
 QY 421 VHODEKQKQKVFQGNFNDNTHRKNVNDPPYARVIRLIPMSWGRITLSELLGCAEE 480
 DB 2074 LFLDQGH--TKVFOGNDSSTPVNALDPPLETRVLRHPRTSMOAHILRLVLGCEAD 2131

RESULT 11
 NRPI_HUMAN STANDARD; PRT; 923 AA.
 AC 014786; 060461;
 DT 30-MAY-2000 (Ref. 39, Created)
 DT 01-OCT-2000 (Ref. 40, Last sequence update)
 DT 01-OCT-2000 (Ref. 40, Last annotation update)
 DE NEUROPILIN-1 PRECURSOR (VASCULAR ENDOTHELIAL CELL GROWTH FACTOR 165 RECEPTOR).
 GN NRPI OR NRP OR VEGF165R.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eulalia; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM).
 RX MEDLINE=97433084; PubMed=9288753;
 RA He Z., Tessier-Lavigne M.;
 RT "Neuropilin is a receptor for the axonal chemorepellent semaphorin 111.";
 RL Cell 90:739-751(1997).
 RN 12
 RP SEQUENCE FROM N.A. (MEMBRANE-BOUND ISOFORM), AND SEQUENCE OF 22-39.
 RC TISSUE-Breast;
 RX MEDLINE=9818809; PubMed=9529250;
 RA Soker S., Takashima S., Miao H.-O., Neufeld G., Klagsbrun M.;
 RT "Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor.";
 RL Cell 92:735-745(1998).
 RN 13
 RP SEQUENCE FROM N.A. (SOLUBLE/SNRPI ISOFORM), AND SEQUENCE OF 22-31.
 RC TISSUE-Prostatic adenocarcinoma;
 RX MEDLINE=20183929; PubMed=10688880;

RA Gagnon M.L., Bieleberg D.R., Gechtman Z., Miao H.-O., Takashima S.,
 RA Soker S., Klagsbrun M.;
 RT "Identification of a natural soluble neuropilin-1 that binds vascular
 RT endothelial growth factor: In vivo expression and antitumor
 RT activity.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:2573-2578(2000).
 RN 14
 RP CHARACTERIZATION.
 RX MEDLINE=20309748; PubMed=10748121;
 RA Guzman-Poltorak Z., Cohen T., Herzog Y., Neufeld G.;
 RT "Neuropilin-2 and neuropilin-1 are receptors for the 165-amino acid
 RT form of vascular endothelial growth factor (VEGF) and of placenta
 RT growth factor-2, but only neuropilin-2 functions as a receptor for
 RT the 145-amino acid form of VEGF.";
 RL J. Biol. Chem. 275:18040-18045(2000).
 CC -1- FUNCTION: THE MEMBRANE-BOUND ISOFORM IS A RECEPTOR INVOLVED IN THE
 CC DEVELOPMENT OF THE CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE
 CC FORMATION OF CERTAIN NEURONAL CIRCUITS AND IN ORGANOPULSANT
 CC OUTSIDE THE NERVOUS SYSTEM. IT MEDIATES THE CHEMOREPULSANT
 CC ACTIVITY OF SEMAPHORINS. IT BINDS TO SEMAPHORIN 3A, THE PLGF-2
 CC ISOFORM OF PGF, THE VEGF-165 ISOFORM OF VEGF AND VEGF-B.
 CC COEXPRESSION WITH KDR RESULTS IN INCREASED VEGF-165 BINDING TO KDR
 CC AS WELL AS INCREASED CHEMOTAXIS. IT MAY REGULATE VEGF-INDUCED
 CC ANGIOGENESIS.
 CC -1- FUNCTION: THE SOLUBLE/SNRPI ISOFORM BINDS VEGF-165 AND APPEARS
 CC TO INHIBIT ITS BINDING TO CELLS. IT MAY ALSO INDUCE APOPTOSIS BY
 CC SEQUESTERING VEGF-165. MAY BIND AS WELL VARIOUS MEMBERS OF THE
 CC SEMAPHORIN FAMILY. ITS EXPRESSION HAS AN AVERSE EFFECT ON BLOOD
 CC VESSEL NUMBER AND INTEGRITY.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE SOLUBLE/SNRPI
 CC ISOFORM IS SECRETED.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; MEMBRANE-BOUND (SHOWN HERE)
 CC AND SOLUBLE/SNRPI; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: THE EXPRESSION OF MEMBRANE-BOUND AND
 CC SOLUBLE/SNRPI ISOFORMS DOES NOT SEEM TO OVERLAP. THE MEMBRANE-
 CC BOUND ISOFORM IS EXPRESSED BY THE BLOOD VESSELS OF DIFFERENT
 CC TISSUES. IN THE DEVELOPING EMBRIO IT IS FOUND PREDOMINANTLY IN THE
 CC NERVOUS SYSTEM. IN ADULT TISSUES, IT IS HIGHLY EXPRESSED IN HEART
 CC AND PLACENTA; MODERATELY IN LUNG, LIVER, SKELETAL MUSCLE, KIDNEY
 CC AND PANCREAS; AND LOW IN ADULT BRAIN. THE SOLUBLE/SNRPI ISOFORM IS
 CC FOUND IN LIVER HEPATOCYTES, KIDNEY DISTAL AND PROXIMAL TUBULES.
 CC -1- SIMILARITY: BELONGS TO THE NEUROPILIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF018956; AAC51759.1; -
 DR EMBL; AF016050; AAC12921.1; -
 DR EMBL; AF145712; AAF44344.1; -
 DR MIM; 602069; -
 DR InterPro; IPR000421; -
 DR InterPro; IPR000859; -
 DR InterPro; IPR000998; -
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00754; F5_F8_TypeC; 2.
 DR Pfam; PF00629; MAM; 1.
 DR PRINTS; PRO0020; MAMDOMAIN.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01285; FAS8C_2; 2.
 DR PROSITE; PS01285; FAS8C_2; 2.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS00740; MAM_2; 1.
 DR PROSITE; PS00060; MAM_2; 1.
 DR Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
 KW Alternative splicing.
 FT SIGNAL 1 21

FT CHAIN 22 923 NEUROFILIN-1.
 FT DOMAIN 22 856 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 857 879 POTENTIAL.
 FT DOMAIN 880 923 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 27 141 CUB 1.
 FT DOMAIN 147 265 CUB 2.
 FT DOMAIN 275 424 F5/8 type C 1.
 FT DOMAIN 431 583 F5/8 type C 2.
 FT DOMAIN 645 811 MAM.
 FT DISULFID 27 54 PROBABLE.
 FT DISULFID 82 104 PROBABLE.
 FT DISULFID 147 173 PROBABLE.
 FT DISULFID 206 228 PROBABLE.
 FT DISULFID 275 424 BY SIMILARITY.
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 842 842 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPPLIC 644 644 EPP -> GIK (IN SOLUBLE/SNRP1 ISOFORM).
 FT VARSPPLIC 645 923 MISSING (IN SOLUBLE/SNRP1 ISOFORM).
 FT CONFLICT 26 26 K -> E (IN REF. 1).
 FT CONFLICT 749 749 D -> H (IN REF. 2).
 FT CONFLICT 853 853 E -> D (IN REF. 2).
 SO SEQUENCE 923 AA: 103120 MW: ADEADC4A849E5D57 CRC64;

Query Match 17.9%; Score 478; DB 1; Length 923;
 Best Local Similarity 31.6%; Pred. No. 8.8e-29;
 Matches 156; Conservative 65; Mismatches 207; Indels 66; Gaps 21;

12 GLGLGPOGKDKDINPCNGGICGLADDSF--SCECEGFRAGNCSSV-EVAD 68
 133 GFIREIREIKRPECSCQNTTPSCVITSPGPEKYNLSLECIYIFAKRMEIILEESF 192
 69 EEKPTAGCICIPNCHNGCTEISEAYRGTF-----IGYVC--KCP--RGNGIHQ 117
 193 DLEPDSNMP-----GGMFCRYDLRIWGFDPVGPPIGRYCGGKTGRIRSSSGILSM 245
 118 HNINECEAPRCNNGICITDLVANSY--ECPCGFMRCQYKCSGHLIEGGIISNOIT 175
 246 --VFYDTSALAKEG-----FSANYSVLQSSVSE-----DFKCMALCMESGEIHSQIT 292
 176 ASSNHRLALGLOKWPYARILNKKGLINMTAENDRMWPIQINLORMKRTGYITGA- 234
 293 ASSQYST-----NMSAERSRLNTE--NGWTPGEDSYREHIOYGLRFTANGTGCAT 345
 235 -KRIGSPETIKSYKIAYNSNDKTYAMTKVKGTEEMVFRGANDVNTPYANSFPPPIAAY 293
 346 SKETKRRKYVYKTKIDVSSNGEDM--TTIKEGNKPVLFGGNTNPPDVVAVFPPPLITRF 403
 294 VRLYPOICRRHCTLMELJGCELSG--CSEPLGMSGHIQDIQITASSVFETLNMDEFTW 351
 404 VRIKPTWETGISMREYVYCKITDPCSGMLGMSGLISDSQITSS-----NQGGRNW 457
 352 EEPKRL--DKQKVNMT-----SGHNDQSOIMLOVLLVTKYTGITITGADPHGVQVYG 406
 458 MPENILVYTRSG--WALPPAPHASTINEMLOIDIGEKEYIAGILOGGHNREKVFMR 513
 407 SYKLAYSNDGEHMYHODEKOROKDVQGNFNDTRKKNVLDPIYARFIRLP-WSWYG 465
 514 KFKIGSYNNGSDMKMIMDSKRAKSEGNNNYDTDELRTF- PALSTFRIRIYEPERATHG 572
 466 RITIRSELGCAEE 479
 573 GLGLRMELLGCEVE 586

RESULT 12
 ID NRPL_XENLA STANDARD: PRT; 928 AA.
 AC P28824;

DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN) (A5 ANTIGEN).
 OS Xenopus laevis (African clawed frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=91373458; PubMed=1908252;
 RA Takagi S., Hirata T., Agata K., Mochii M., Eguchi G., Fujisawa H.;
 RT "The A5 antigen, a candidate for the neuronal recognition molecule,
 has homologies to complement components and coagulation factors.";
 RL Neuron 7:295-307(1991).
 CC -1- FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
 CC CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
 CC CERTAIN NEURONAL CIRCUITS AND IN ORGANORENOSTHESIS OUTSIDE THE NERVOUS
 CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (BY
 CC SIMILARITY). PRESUMED TO BE INVOLVED IN THE NEURONAL RECOGNITION
 CC BETWEEN THE OPTIC NERVE FIBERS AND THE VISUAL CENTERS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: RETINAL GANGLION CELLS AND VISUAL CENTER
 CC NEURONS.
 CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10467; BAA01260.1; -
 DR InterPro: IPR000421; -
 DR InterPro: IPR000959; -
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR Pfam: PF00629; MAM; 1.
 DR PRINTS: PR00020; MAMDOMAIN.
 DR PROSITE: PS00740; MAM_1; 1.
 DR PROSITE: PS01180; CUB; 2.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 DR PROSITE: PS00600; MAM_2; 1.
 DR Transmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
 KW Antigen.
 KM
 FT CHAIN 1 21
 FT DOMAIN 22 928
 FT TRANSMEM 861 883
 FT DOMAIN 864 928
 FT DOMAIN 27 141
 FT DOMAIN 147 265
 FT DOMAIN 275 424
 FT DOMAIN 431 584
 FT DOMAIN 646 812
 FT DISULFID 27 54
 FT DISULFID 82 104
 FT DISULFID 147 173
 FT DISULFID 206 228
 FT DISULFID 275 424
 FT DISULFID 431 584
 FT CARBOHYD 261 261
 FT CARBOHYD 300 300

FT CARBOHYD 523 523 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 928 AA; 103416 MW; AF6B323B0A4C789D CRC64;

Query Match 17.7% Score 472.5; DB 1; Length 928;
 Best Local Similarity 31.4%; Pred. No. 2,3e-28;
 Matches 154; Conservative 65; Mismatches 213; Indels 59; Gaps 21;

QY 12 GLSLGVPGFGKGDICNPNCENGICLSGLADDSF--SCCEPGEFAGPNCSSVY-EVAD 68
 DB 133 GFSIRREVFKRPECSSRNFTSSNGVKSPPRYPEKYPALAECTYIIIFAPKMOETVLEFESEF 192
 QY 69 EEKPTSGACPICPNCHNGTCE-----ISEAYNG-DTEIGYVC--KCP---NGFNGICQ 117
 DB 193 ELRADSNAP-----GGOTCRYDMLGIMDGFPGVGHISGRYCGQNTPGRVRSFTGISM 245
 QY 118 HNINECEAPCRNGGICITDLVANYSCCEPGFPMGRNC--QYKCSGHLGIEGIIISNOQIT 175
 DB 246 --IFHDSALAKEG-----FRANFS-----VQOSNIDEPFOCKEALGESEIHFQDIS 292
 QY 176 ASSNHALGLGLOKWPYIARLNKGLINAMTAENDRWPMIQLINQKRMVTVTQGA- 234
 DB 293 VSSQYS-----MMSAERSRLNY--VENGMTPEGDTVKEMIOVDLENLEFVSIGITQAI 345
 QY 235 -KRIGSEYIKSYKIAVNDKRTWAMYKVGNTNEMVFRGNVDNNTPVANSFTPEPKAOY 293
 DB 346 SKETKKRYFVKSYKVDISSNGEDM--ITLKDGNKHLVFTGNTDPTDVPYRPRSKFVITRF 403
 QY 294 VRIYPOICRRHCTLRNELLGELSG--CSEPLKRSKHIOVDVOTFSSVFRTLNMDMTW 351
 DB 404 VRLRPYTWNGISLRFELYGCKITIDYPCSRMLGVSLDISQITASS-----QYDR-NM 457
 QY 352 EEPKATL--DKOGKYNATSGHNDOSOMLOYDLVPRKTYGIIQTQAKRFGVQVGSXK 409
 DB 458 VPLAELVSRSGMALPSPNTHPYTKENLODLAEKRYAGVYIGQKKEKRYMKRFK 517
 QY 410 LAYSNDGEMHWVHODEKQRKDVFOGNDNDRKNVYIDPIYARITRLP-WSYGRIT 468
 DB 518 IGSNNGTWEMIMDSKNKPKTEGNTYDTELRTF-AHITTGRIIRPERASAGSLA 576
 QY 469 LRSELLGCAEE 479
 DB 577 LRLELGCVE 587

RESULT 13
 NRPL-CHICK
 ID NRPL-CHICK STANDARD: PRT: 914 AA.
 AC P79755;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROFILIN-1 PRECURSOR (A5 PROTEIN).
 GN NRPI OR NRP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE LEHORN; TISSUE-Embryonic brain;
 RA MDLLine-95324761; PubMed-7601310;
 RA Takai S., Kasuya Y., Shimizu M., Matsuura T., Tsuboi M., Kawakami A.,
 RA Fujisawa H.;
 RT Expression of a cell adhesion molecule, neuropilin, in the
 RT developing chick nervous system.";
 RT Dev. Biol. 170:207-222(1995).
 RL
 CC FUNCTION: RECEPTOR INVOLVED IN THE DEVELOPMENT OF THE
 CC CARDIOVASCULAR SYSTEM, IN ANGIOGENESIS, IN THE FORMATION OF
 CC CERTAIN NEURONAL CIRCUITS AND IN ORGANOGENESIS OUTSIDE THE NERVOUS
 CC SYSTEM. IT MEDIATES THE CHEMOREPULSANT ACTIVITY OF SEMAPHORINS (B1

CC SIMILARITY. SEEMS TO HAVE CALCIUM-INDEPENDENT CELL ADHESION
 CC PROPERTIES.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: DEVELOPING NERVOUS SYSTEM; OPTIC TECTUM
 CC (LAYERS D AND E OF SGFS), AMACRINE CELLS OF RETINA, NEURITES OF
 CC DORSAL ROOT GANGLIA. ALSO EXPRESSED IN NONNEURONAL CELLS, E.G.
 CC BLOOD VESSELS IN THE ENTIRE EMBRYO.
 CC -1- SIMILARITY: BELONGS TO THE NEUROFILIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D45416; BA08256.1; -;
 CC InterPro: IPR000421; -;
 CC InterPro: IPR000859; -;
 CC InterPro: IPR000998; -;
 CC Pfam: PF00431; CUB; 2.
 CC Pfam: PF00754; F5_F8_type_C; 2.
 CC Pfam: PF00629; MAM; 1.
 CC PRINTS: PR00020; MAMDOMAIN.
 CC PROSITE: PS01180; CUB; 2.
 CC PROSITE: PS01285; FAS8C.1; 2.
 CC PROSITE: PS01286; FAS8C.2; 2.
 CC PROSITE: PS00740; MAM.1; 1.
 CC PROSITE: PS5060; MAM.2; 1.
 CC Trnsmembrane; Glycoprotein; Neurone; Signal; Repeat; Receptor;
 KW Cell adhesion.
 KW SIGNAL 1 18
 FT CHAIN 19 914
 FT DOMAIN 20 847
 FT TRANSSEM 848 870
 FT DOMAIN 871 914
 FT DOMAIN 25 139
 FT DOMAIN 145 263
 FT DOMAIN 273 422
 FT DOMAIN 429 581
 FT DOMAIN 636 801
 FT DOMAIN 25 52
 FT DISULFID 80 102
 FT DISULFID 145 171
 FT DISULFID 204 226
 FT DISULFID 273 422
 FT DISULFID 429 581
 FT SEQUENCE 914 AA; 102480 MW; DD2EB6DF0CB868C CRC64;

Query Match 17.5% Score 468; DB 1; Length 914;
 Best Local Similarity 30.5%; Pred. No. 5e-28;
 Matches 152; Conservative 70; Mismatches 201; Indels 76; Gaps 21;

QY 12 GLSLGVPGFGKGDICNPNCENGICLSGLADDSF--SCCEPGEFAGPNCSSVY-EVAD 68
 DB 133 GFSIRREVFKRPECSSRNFTSSNGVKSPPRYPEKYPALAECTYIIIFAPKMOETVLEFESEF 190
 QY 69 EEKPTSGACPICPNCHNGTCEISEAYNGDTF-----IGVCKCPRGFNGICHQININE 122
 DB 191 ELRPDSWTP-----GGAFCRYDRLEIMDGFPPDVGHIGRYCG-----QNNPGR 233
 QY 123 CEAPPCRNNGIC-----TD-----LVANSCCEPGFPMGRNCQYKSGHLGIEGIGIS 170
 DB 234 VRS-----STILSMVFTYDTSALAKGFSANVSQSSV---SDDPQCKMPLGMEGGEIH 285
 QY 171 NQQTASSNRALFGLOKWPYIARLNKGLINAMTAENDRWPMIQLINQKRMVTVTQGI 230
 DB 286 SDQITVSSQYSAL-----WSSERSRLNTP--NGWTPEDBSVREMIOVDLGLLRVSGIG 338

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:26 ; Search time 105.14 Seconds
(without alignments)
535.094 Million cell updates/sec

Title: US-09-237-981-10
Perfect score: 2669
Sequence: 1 MKHLVAMLVGLVSLGVPPF.....MSWYGRITLRSLLGCAEE 480

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SPTRMBL_15:*
 - 2: sp.bacteria:*
 - 3: sp.fungi:*
 - 4: sp.human:*
 - 5: sp.invertebrate:*
 - 6: sp.mammal:*
 - 7: sp.mhc:*
 - 8: sp.organelle:*
 - 9: sp.phage:*
 - 10: sp.plant:*
 - 11: sp rodent:*
 - 12: sp.unclassified:*
 - 13: sp.vertebrate:*
 - 14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2649	99.3	480	11	035474
2	2559	95.9	480	11	043854
3	1189.5	44.6	463	11	09R1X9
4	1169	43.8	426	11	09WT53
5	987.5	37.0	363	6	077718
6	694	26.0	2183	11	088783
7	684	25.6	2224	4	043737
8	650.5	24.4	2343	6	062730
9	646.5	24.2	2343	6	018806
10	493.5	18.5	858	5	076470
11	462.5	17.3	216	4	014286
12	462	17.3	779	4	09NRT3
13	460	17.2	921	11	09GX38
14	436	16.3	192	11	09QWQ1
15	341	12.8	2352	5	061240
16	332	12.4	406	5	025059
17	324	12.1	3644	5	09VU94
18	324	12.1	3843	5	09USD0
19	317	11.9	721	13	091902

20	314	11.8	728	13	090656	Q90656 gallus gall
21	313	11.7	529	5	025058	Q25058 helioicdarii
22	312.5	11.7	2653	5	025253	Q25253 lucilia cup
23	311	11.7	2447	13	013149	Q13149 fuqua rubrip
24	308.5	11.6	2634	5	09W4T8	Q9W4T8 drosophila
25	308.5	11.6	2704	5	097458	Q97458 drosophila
26	307.5	11.5	717	13	P87357	P87357 brachydantio
27	303	11.4	752	13	Q42374	Q42374 brachydantio
28	303	11.4	1203	11	006008	006008 mus musculu
29	303	11.4	2470	11	035516	035516 mus musculu
30	301.5	11.3	723	4	Q9N041	Q9N041 homo sapien
31	301	11.3	585	5	09U0E2	Q9U0E2 tribolium c
32	299.5	11.2	2531	5	016004	016004 lytechinus
33	298.5	11.2	802	13	057462	057462 brachydantio
34	298	11.2	723	4	Q9UJY2	Q9UJY2 homo sapien
35	297.5	11.1	615	13	057409	057409 brachydantio
36	297	11.1	1404	5	Q9VB65	Q9VB65 drosophila
37	297	11.1	2471	11	Q9QW30	Q9QW30 rattus sp.
38	293.5	11.0	2319	11	Q9R172	Q9R172 rattus norv
39	292.5	11.0	1530	11	Q9WUG5	Q9WUG5 rattus norv
40	292	10.9	224	11	Q9R1M6	Q9R1M6 mus musculu
41	292	10.9	529	5	Q9V048	Q9V048 drosophila
42	292	10.9	1193	13	Q90819	Q90819 gallus gall
43	291	10.9	1218	4	015122	015122 homo sapien
44	291	10.9	1218	4	015816	015816 homo sapien
45	291	10.9	1218	4	014902	014902 homo sapien

ALIGNMENTS

RESULT 1
ID 035474 PRELIMINARY: PRT: 480 AA.

AC 035474: 035475:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE INTEGRIN-BINDING PROTEIN DELTA PRECURSOR.
GN EDI13 OR DEL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC TISSUE=EMBRYO;
RX MEDLINE=96083109; PubMed=9420328;
RA Hidayi C., Zupancic T., Penta K., Mikhail A., Kawana M.,
RA Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
RA Auerbach R., Hogan B.L.M., Snodgrass R., Quettermous T.;
RT "Cloning and characterization of developmental endothelial locus-1: an
RT embryonic endothelial cell protein that binds the alphabeta3 integrin
RT receptor".
RT Genes Dev. 12:21-33(1998).
CC - FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
CC FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
CC REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC
CC DEVELOPMENT.
CC - SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC - ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND
CC SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
CC CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RESTRICTED GROUP
CC OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
CC NEURONS.
CC - DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7, AFTER DAY
CC 15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME
CC OF BIRTH.
CC - SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC - SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR EMBL: AF031524; AAB86585.1; -.

Query Match	Similarity	99.3%	Score	2649	DB	11	Length	480
Best Local	Similarity	99.4%	Pred.	No.4,1e-229				
Matches	477	Conservative	1	Mismatches	2	Indels	0	Gaps
QY	1	MKHLVAMLLVGLSLGVPQFGKGDICNPENPCENGICLSGLADDSFSCDECPDFAGPNC	60					
Db	1	MKHLVAMLLVGLSLGVPQFGKGDICNPENPCENGICLSGLADDSFSCDECPDFAGPNC	60					
QY	61	SVYEVASDEEKPTSPAGPCIPNCHNGGCEISEAYRGDTPIGYCKCKPRPFGNHCQHN	120					
Db	61	SVYEVASDEEKPTSPAGPCIPNCHNGGCEISEAYRGDTPIGYCKCKPRPFGNHCQHN	120					
QY	121	NECEAEPCRNCGICTDVLVANYSCDEPGFPMRNCQYKCSGHLIEGGLISNQGITASSNH	180					
Db	121	NECEAEPCRNCGICTDVLVANYSCDEPGFPMRNCQYKCSGHLIEGGLISNQGITASSNH	180					
QY	181	RALEFLOKKYPPYARLNKKKGLINMTAAENDRWPMIOINLORKKRYGVITOGAKRTGSP	240					
Db	181	RALEFLOKKYPPYARLNKKKGLINMTAAENDRWPMIOINLORKKRYGVITOGAKRTGSP	240					
QY	241	EYIKSYKIAVNSDGTWMAVKVGTNEEMVFRGVNDNNTPYANSFPPIKAQVRLYPTI	300					
Db	241	EYIKSYKIAVNSDGTWMAVKVGTNEEMVFRGVNDNNTPYANSFPPIKAQVRLYPTI	300					
QY	301	CRRHCTLRRELLGCELSGCEPGLMKSGHIDYQITASSVFRLLNMDFTWBRKARLDK	360					
Db	301	CRRHCTLRRELLGCELSGCEPGLMKSGHIDYQITASSVFRLLNMDFTWBRKARLDK	360					
QY	361	OGKVAAMTISGSHDQSQMLQVLDLVPFKTGTGITTGADGDFHVOFGVSYKLAANSNGEHHM	420					
Db	361	OGKVAAMTISGSHDQSQMLQVLDLVPFKTGTGITTGADGDFHVOFGVSYKLAANSNGEHHM	420					
QY	420	OGKVAAMTISGSHDQSQMLQVLDLVPFKTGTGITTGADGDFHVOFGVSYKLAANSNGEHHM	480					
Db	420	OGKVAAMTISGSHDQSQMLQVLDLVPFKTGTGITTGADGDFHVOFGVSYKLAANSNGEHHM	480					

Db	421	VHODEKORDKVQGCNFNDNRKKNVIDPPIYAFINILPMSVNGRTTLRSELLGCAEE	480
RESULT	2		
ID	043854	PRELIMINARY;	PRT; 480 AA.
AC	043854; 043855;		
DT	01-JUN-1998 (TREMblrel.06, Created)		
DT	01-JUN-1998 (TREMblrel.06, Last sequence update)		
DT	01-OCT-2000 (TREMblrel.15, Last annotation update)		
DE	INTEGRIN-BINDING PROTEIN DELTA PRECURSOR.		
GN	DELTA.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.		
RC	TISSUE-EMBRYONIC LUNG;		
RX	MEDLINE=98083109; PubMed=9420328;		
RA	Hidal C., Zupancic T.J., Penta K., Mikhail A., Kawana M.,		
RA	Quettermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platina D.,		
RA	Auerbach R., Hogan B.L.M., Snodgrass R., Quettermous T.;		
RT	"Cloning and characterization of developmental endothelial locus-1: an		
RT	embryonic endothelial cell protein that binds the alphavbeta3 integrin		
RT	receptor.";		
RL	Genes Dev. 12:21-33(1998).		
CC	-1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH		
CC	INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS		
CC	FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN		
CC	REGULATION OF VASCULAR MORPHOGENESIS OF REMODELING IN EMBRYONIC		
CC	DEVELOPMENT.		
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.		
CC	-1- ALTERNATIVE PRODUCTS: TWO ISOFORMS (SHOWN HERE) AND		
CC	SHORT ISOFORM: ARE PRODUCED BY ALTERNATIVE SPLICING.		
CC	-1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.		
CC	-1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.		
DR	EMBL; U70312; AAC02648.1; -		
DR	HSSD; P00740; I1XA.		
DR	INTERPRO: IPR000152; -		
DR	INTERPRO: IPR000421; -		
DR	INTERPRO: IPR000561; -		
DR	INTERPRO: IPR000742; -		
DR	INTERPRO: IPR001881; -		
DR	PFAM; PF00008; EGF_3.		
DR	PFAM; PF00754; F5_F8_Type_C_2.		
DR	PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.		
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.		
DR	PROSITE; PS01186; EGF_2; UNKNOWN_2.		
DR	PROSITE; PS01187; EGF_CA_1.		
DR	PROSITE; PS01285; FAS8C_1; 2.		
DR	PROSITE; PS01286; FAS8C_2; 2.		
KW	EGF-like domain; Alternative splicing; Signal; Developmental protein;		
KW	Cell adhesion; Repeat; Vascularization.		
FT	SIGNAL	1	16
FT	CHAIN	17	480
FT	DOMAIN	26	59
FT	DOMAIN	78	116
FT	DOMAIN	123	154
FT	DOMAIN	161	311
FT	DOMAIN	322	473
FT	SITE	96	98
FT	DISULFID	26	37
FT	DISULFID	31	48
FT	DISULFID	50	59
FT	DISULFID	78	89
FT	DISULFID	83	105
FT	DISULFID	107	116
FT	DISULFID	158	314
FT	DISULFID	301	305
FT	DISULFID	319	476
FT	DISULFID	319	476

Wed May 23 06:29:16 2001

us-09-237-981-10.rsp

Page 4

DR	PROSITE; EGF_2; 2.	
DR	PROSITE; PS01285; FA58C_1; 2.	
DR	PROSITE; PS01286; FA58C_2; 2.	
SQ	SEQUENCE 426 AA; 47197 MW; B182F4A206298B1A CRC64	

Query Match	43.8%;	Score 1169;	DB 11;	length 426;
Best Local Similarity	48.2%;	Pred. No. 1.4e-96;		
Matches 220;	Conservative 69;	Mismatches 111;	Indels 56;	Gaps 7;

```

OY 23 GDIJGNPNBENCJOLGSLADSPSCGCEPFGPNCSSVVEVADDEKPSIAPRCIPNP 82
Db 25 GPCDSJSLCJNGJCTLU-QDNIDYICLCPBGTGLVONE-----TERGCSNP 72
OY 83 CHNGJTCETS-EAYRGDTFJGYVCKCPRGFNHICOHINICEABPCRNNGICTDLVANY 141
Db 73 CYNDAJCETLDTLDTGDTJFETEXICQCCVGVSGJHCTGT----- 110
OY 142 SCECPGPEFMRNRCOKYCSGSHGJGIGGJISMOQJITASNNHRLFGLOKNYPYARLANKGL 201
Db 111-----GSTOLGMEGALMADSOIASASVYMGMFGDORMGPELARLYRNGI 154
OY 202 INAMTAENDRMPIOMJLOKRMKRVGVTJOGAKRIGSEPEYIKSVIKIYXSDKCTAMAYK 261
Db 155 VAMTAISNTDSKPMJQVWLLKMKMVSVMYOGASRBRGAEVLTKTKVMSIDSKFEFIQ 214
OY 262 -VKGTEENKVERGVNNTFPYANSFTPIKAQTVRLYPOICRRHCTLRLMELLCSELSGS 320
Db 215 DESGGKRE-FGJNIDNNJSLVNNFNFTLEAOJIRLYPVSCRGCTJREFSELGCELGCS 272
OY 321 EPJGAKMSGJIDJYQJITASSVFRJLNMDFMFWERBRARALDOKKVAWANSCHNDOSOMLOV 380
Db 273 EPJGLKNNJIPDSOMASASSKYTNLPAFCWYBHLGRJLDNOKKINAMTAQNSAKEMWLOV 332
OY 381 DCLVPRKVGJITOGAKRFGVGVFSYKLAAYNDEEHMVAJODKOKDKVPOGNGND 440
Db 333 DCLGQKQVJGJITOGARBFGJIOVYASIKYKAHSDDGVOMTYV--EEOGSSKVFQNGNDNN 390
OY 441 THKRVNIDPPIYARJIRLLPMSTGRTTLESELGC 476
Db 391 SHKNNFEKPEMARVVRVLVSMNNRJTLYTLELGC 426

```

RESULT	5			
077718				
ID	077718	PRELIMINARY:	PRT:	363 AA.
AC	077718			
DT	01-NOV-1998 (TREMBL)	08, (Created)		
DT	01-NOV-1998 (TREMBL)	08, (Last sequence update)		
DT	01-MAY-2000 (TREMBL)	13, (Last annotation update)		
DE	SPERM-MEMBRANE ASSOCIATED PROTEIN P47 (FRAGMENT).			
OS	Equus caballus (Horse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eutheria; Perissodactyla; Equidae; Equus.			
NCBI_TaxID	9796;			
NCBI_TaxID	9796;			
SEQUENCE	FROM N.A.			
TISSUE	TESTIS:			
RA	Genzel M., Toepfer-Petersen E.,			
RA	Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases			
DR	EMBL, Acc10121; CA09010.1; -.			
DR	INTERPRO: IPR000421; -.			
DR	INTERPRO: IPR000561; -.			
DR	PFAM: PF00008; EGF_1.			
DR	PFAM: PF00754; F5_F8_type_C_2.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.			
DR	PROSITE: PS01186; EGF_2; 2.			
DR	PROSITE: PS01285; FA8C_1; 1.			
DR	Sperm.			
KW	NON_TER			
FT	1			
FT	363			
FT	363			
SEQUENCE	363 AA; 40744 MW; 1F8B6395AF32338D CRC64;			

Query Match	37.0%;	Score 987.5;	DB 6;	Length 363;
Best Local Similarity	43.1%;	Pred. No. 2e-60;		
Matches 187;	Conservative 67;	Mismatches 107;	Indels 73;	Gaps 5;

[illegible]

RESULT	6		
008783			
ID	008783	PRELIMINARY;	PRT; 2183 AA.
AC	008783;		
DT	01-NOV-1998 (TREMBLrel. 08, Created)		
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)		
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)		
DE	MURINE COAGULATION FACTOR V.		
GN	F5.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_Taxid=10090;		
NP	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yang T.L., Cui J., Rehmuculla A., Mousallil M., Kaufman R.J.,		
RA	Ginsburg D.;		
RT	"The structure and function of murine factor V and its inactivation by		
RL	protein C.";		
RL	Blood 91:0-0(0012).		
KN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Ginsburg D., Yang T.L., Cui J., Yang A.;		
RL	Submitted (Mar-1996) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; 052925; AAC09553.1; .		
DR	HSSP; P00450; IKCW.		
MGI	MGI:88382; F5.		
DR	INTERPRO: IPR000421; .		
DR	INTERPRO: IPR01117; .		
DR	PFAM: PF00394; Cu-oxidase; 3.		
DR	PFAM: PF00754; F5_E8_type.C.2.		
DR	PROSITE: PS00079; MULTICOOPER_OXIDASE1; 2.		
DR	PROSITE: PS01285; FA36C_1; 2.		
DR	PROSITE: PS01286; FA36C_2; 2.		
Q0	SEQUENCE 2183 AA; 247328 MW; BF0A8AA723FE60317 CRC64;		

Query Match 26.0%; Score 694; DB 11; Length 2183;
Best Local Similarity 43.8%; Pred. No. 4,4e-53;
Matches 144; Conservative 51; Mismatches 112; Indels 22; Gaps 5;

157 KCSGHLGTEGGIISNOQTASSNHRALFGLQKWPYARLNKKGLINAW---TAENDR 212
1865 ECKMKGSLSTGYISDSQIKASEY-----LYWEPRLARLNAGSYNAMSLEKTRALDPPI 1918
213 WPMIOLNLRKMRVGVITOGAKRIGSPYIKS-----YKLAISNDGKTWAMYKKGTE 267
1919 KPMIOVDMQKEVYVIGILOGAK-----HYKSCFTTEFOVAYSDDQNMWOLFEGKSGKS 1973
268 EAVFPGNDNNTPYANSFPPIKAOYVRLYPOICRRHCTLRMELLCGELSGSEPLGKMS 327
1974 VMFTGNSDGTSTIKENRDLDPPIVARYIRIHPKSYNRPRLRELOGCEVNGCSTPLGED 2033
328 GHIDVQITASSVFTLNMDMTPEPRKARLDKQGVANWTSGHNDQSQWLOVDLLVPTK 387
2034 GRIDQKQITASSFKKSMWDY--WEPSLARLNAGGVANWAKANNKQWLOVDLLKTKK 2091
388 VTGIIYOGAKDFGHVOFVGSYKLAYSNDGEHMHVHODEKORKDKVFOGNDTHRKVNI 447
2092 VTAIVTQCGCKSLSEMYVKSYSIQYSDQGVAMKPYRQKSSMYDKLFEQNSNTKGMKNPF 2151
448 DPEIYARFIRILPMSWYGRITRLSELGCG 476
2152 NPEIISRFIRIIPKTMNOSIALRLLEFGC 2180

RESULT 7
043737 PRELIMINARY; PRT: 2224 AA.

ID 043737;
AC 043737;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE FACTOR V.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; 299572; CAB16748.1; -
DR HSSP; P00450; ICKW.
DR INTERPRO: IPR000421; -
DR INTERPRO: IPR001117; -
DR PFAM; PF00394; Cu-oxidase; 3.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
KM Blood coagulation.
SQ SEQUENCE 2224 AA; 251673 MW; 1013108D49690EAB CRC64;

Query Match 25.6%; Score 684; DB 4; Length 2224;
Best Local Similarity 42.4%; Pred. No. 3.5e-52;
Matches 143; Conservative 55; Mismatches 123; Indels 16; Gaps 4;

144 ECPGEFMRNCOYKCSGHLGIEGGIISNOQTASSNHRALFGLQKWPYARLNKKGLIN 203
1897 QTFPLIMDRCCRP---MGLSTGIISDSQIKASE-----FLGYWEPRLARLNAGSYN 1946
204 AAF-----AANDRWPMIOLNLRKMRVGVITOGAKRIGSPYIKSYKLAYSNDGKTWAM 259
1947 AMSVEKLAFAFKPMTQVDMQKEVITIGIOTGAKHNYLSCYTTTFYVAISSNOIWMQI 2006
260 YKVGTEMEVYRGVNDNNTPYANSFTPIKAOYVRLYPOICRRHCTLRMELLCGELSGC 319

2007 FKGNSTRNWMYFNGNSASTIKENQDPPIVARYIRISPRAYNRPLRLRELOGCEVNGC 2066
320 SEPLGKSGHIQDYQITASSVFTLNMDMTPEPRKARLDKQGVANWTSGHNDQSQWLO 379
2067 STPLGMEKNGIENKQITASSFKKSMWDY--WEPSLARLNAGGVANWAKANNKQWLOE 2124
380 VDLVPTKVTGIIYOGAKDFGHVOFVGSYKLAYSNDGEHMHVHODEKORKDKVFOGNDN 439
2125 IDLLKIKRITAIITQCGCKSLSEMYVKSYSITIHSEQVEKMPYRLKSSMYDKLFEQNTWT 2184
440 DTHRKNVIDPEIYARFIRILPMSWYGRITRLSELGCG 476
2185 KGHVKNFNPPIISRFIRIIPKTMNOSIALRLLEFGC 2221

RESULT 8
062730 PRELIMINARY; PRT: 2343 AA.

ID 062730;
AC 062730;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE FACTOR VIII.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-KIDNEY, SPLEEN;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF049489; AAC05384.1; -
DR HSSP; P00451; ICFG.
DR INTERPRO: IPR000421; -
DR INTERPRO: IPR001117; -
DR PFAM; PF00394; Cu-oxidase; 3.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744DAAD99 CRC64;

Query Match 24.4%; Score 650.5; DB 6; Length 2343;
Best Local Similarity 43.1%; Pred. NO. 3.0e-49;
Matches 140; Conservative 51; Mismatches 119; Indels 15; Gaps 5;

157 KCSGHLGTEGGIISNOQTASSNHRALFGLQKWPYARLNKKGLINAWTAENDRWPMI 216
2031 KCQPLGMAASHIRDFQITASSGY-----GQWPKLARLHLYSSINAMST--KDPFSWI 2082
217 QINLORKMRVGVITOGAKRIGSPYIKSYKLAYSNDGKTWAMYKVGTEMEVYRGVND 276
2083 KYDLAPMIGHIMOGARQKFFSLYSQFLIMYSLDGKMKHWSRGNSTGTLAVFQNV 2142
277 NNTPYANSFTPIKAOYVRLYPOICRRHCTLRMELLCGELSGSEPLGMSGHIQDYQIT 336
2143 SSGIKHNLFNPPITAOYVRLYRHLPTHYSIRSTLRMELLCGDFNSCSPMGESKALSDAQIT 2202
337 ASSVPTLNMDMTPEPRKARLDKQGVANWTSGHNDQSQWLOVDLLVPTKVTGIIYOGA 396
2203 ASSYISSM---LATPSPQALHLHLOGRINARNRPPANNKEKLOYDFKTKMYIGITGGV 2259
397 KDFGHVOFVGSYKLAYSNDGEHMHVHODEKORKDKVFOGNDTHRKVNIPEIYARF 455
2260 KSLTISMVYKFEFLISSDQGNWTLFLQNDKV---KVFQGRDSDSTVRNALPEPIVARY 2316
456 IRIIPMSWYGRITRLSELGCAEE 480
2317 VRLHPOSNAHHLALRLLEVLGCDTQO 2341

```

RESULT 9
ID 018806 PRELIMINARY; PRT; 2343 AA.
AC 018806:
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE FACTOR VIII.
GN F8.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Cameron C., Nolley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
RA Gilles A., Lilligrip D.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016234; AAB87412.1; -.
DR HSSP; P00451; ICFG.
DR INTERPRO; IPR000421; -.
DR INTERPRO; IPR001117; -.
DR PRAM; PF00394; Cu-oxidase; 3.
DR PFAM; PF00754; F5_F8_type-C; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;

Query Match 24.2%; Score 646.5; DB 6; Length 2343;
Best Local Similarity 42.9%; Pred. No. 8.7e-49;
Matches 139; Conservative 50; Mismatches 122; Indels 13; Gaps 4;

OY 157 KCSGHIGEGGIISSNOOITASSNHRALFGLQKRYPYARLNKGLINATAENDRWPI 216
DB 2031 KCTPLGMSGHRFDPOITASGOY-----GWAFLRLHYSGSINAST--KDFSWI 2082
OY 217 QINLQKRMVYGVITGAKRIGSPYIKSYKIAYSDGKTWAMYKKGTEMEVFRGND 276
DB 2083 KYDLAPMIHIGITGAGRKSSLYVSQFIIMSLDGKMKWISYRNSGTLMVFFGND 2142
OY 277 NMTPLANSFPPIKAQYVRLYPOICRRHCTLMELLIGCELSCESEPLGKSGHIQDYOT 336
DB 2143 SSGIKHNINPRLIAQYIRLHPHYISIRSTLMELLIGCELSCESEPLGKSKAISDAQT 2202
OY 337 ASSVFRLNMDMTWEPKARLDKQKVNAWTSQNDOSQMLQVLDLPTKVTGIIITGCA 396
DB 2203 ASSYLESM---LATWSPQARLHLQRTVAMRPQANNPKEMQLQVDRKTKVGTITGCV 2259
OY 397 KQEGHVOFGYSKYLAYSNDGEHMMVHODEKQKDKVFGQNFNDTHRKNVIDPPIYAREI 456
DB 2260 KLLISMVYKEFLISSQGHMMWTLFL--QNGKVKVFGQNRDSSIPVRRLRPLVARYV 2317
OY 457 RILPWSVGRITLRSLLCAEE 480
DB 2318 RLHPQSMANHIALRLVLCDDTQ 2341

RESULT 10
ID 076470 PRELIMINARY; PRT; 858 AA.
AC 076470:
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE ECHINONECTIN.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.

```

```

OX NCBI_TaxID=7654;
RN 11
RP SEQUENCE FROM N.A.
RA Allegro M.C., Allegro M.A.;
RT "Echinonectin."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF070482; AAC32598.2; -.
DR HSSP; P00451; ICFG.
DR INTERPRO; IPR000421; -.
DR INTERPRO; IPR001092; -.
DR PFAM; PF00754; F5_F8_type-C; 4.
DR PROSITE; PS01285; FA58C_1; 4.
SQ SEQUENCE 858 AA; 94845 MW; 56E43AB727ED848 CRC64;

Query Match 18.5%; Score 493.5; DB 5; Length 858;
Best Local Similarity 38.1%; Pred. No. 1.2e-35;
Matches 127; Conservative 48; Mismatches 117; Indels 41; Gaps 10;

OY 160 GHIGEGGIISSNOOITASSNHRALFGLQKRYPYARLNKGLINATAENDRWPI 212
DB 381 KGLGIEDSISTRISSASS-----CWDSDHCVDNRSLNQPEGSLRGAMSAITNNE 431
OY 213 WPIQINLQKRMVYGVITGAKRIGSPYIKSYKIAYSDGKTWAMYKKGTEMEVFR 272
DB 432 NQMIQVDDLAFAEYGVITG--RNSGNQWESYGISYIDKQWILVD--DCDREPKTFV 488
OY 273 GNVNDNTPYANSFPPIKAQYVRLYPOICRRHCTLMELLG---CELSGCEPLGKSGH 329
DB 489 ANFSDSLVENGISPPYARREFRHLPTWNNHISLRWELLGEGPPTLADASHIIGLDYR 548
OY 330 IQDYQITASSVFRLNMDMTWEPKARLD-----KQKVNAWTSQNDOSQMLQVLD 382
DB 549 IPDGSITSTEDS-----NHAARARLNLPLGALKG---WSASTLDHSGMLQVLD 598
OY 383 LVPKVTGIIITGAKDFGHVQFVGSYKIAYSDGEHMMVHODEKQKDKVFGQNFNDTH 442
DB 599 RGYRVTVGVITGQRSDAN--QWVTEYKVAHSLNGLIEDTJIAAASQDKVFGNSDRSTQ 656
OY 443 RKNVIDPPIYARFRLPWSVGRITLRSLELG 475
DB 657 VTNVFPPLITRFLRILPTMRYGHISMELLG 689

RESULT 11
ID 014286 PRELIMINARY; PRT; 216 AA.
AC 014286:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-2000 (Tremblrel. 13, Last annotation update)
DE COAGULATION FACTOR VIII ASSOCIATED PROTEIN B.
GN F8B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RL MEDLINE=93052386; PubMed=1427887;
RA Leyinson B., Kenwick S., Gamel P., Fisher K., Gitschier J.;
RL "Evidence for a third transcript from the human factor VIII gene."
DR EMBL; M90707; AAB58466.1; -.
DR HSSP; P00451; ICFG.
DR INTERPRO; IPR000421; -.
DR PRAM; PF00754; F5_F8_type-C; 1.
DR PROSITE; PS01285; FA58C_1; UNKNOWN_1.
DR PROSITE; PS01286; FA58C_2; 2.
SQ SEQUENCE 216 AA; 24641 MW; 6C82D4F89E35A376 CRC64;

Query Match 17.3%; Score 462.5; DB 4; Length 216;

```


Db 513 KFKIAYNNNGSDMKIMDDSKRAKSPGEGNNNDTPELRAE-TPLSRFRIRIPEARATIS 571
 QY 466 RITRSELLCAEE 479
 Db 572 GLGIRMELLCVEE 585

RESULT 14

090M01 PRELIMINARY; PRT; 192 AA.

AC 090M01; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE MILK FAT GLOBULE GLYCOPROTEIN (FRAGMENT).
 MG-E8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
 RT "Lactation-dependent expression of an mRNA splice variant with an exon
 for a multiply O-glycosylated domain of mouse milk fat globule
 glycoprotein MG-E8";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB021647; BAA36214.1; JOINED.
 DR INTERPRO; IPR000421; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR001092; -;
 DR PFAM; PF00008; EGF_1;
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS01285; FAS8C_1; 1.
 FT NON_TER 1 192
 FT NON_TER 192
 SQ SEQUENCE 192 AA; 21134 MW; EE110C03A89689F4 CRC64;

Query Match

Best Local Similarity 16.3%; Score 436; DB 11; Length 192;

Matches 83; Conservative 32; Mismatches 61; Indels 6; Gaps 3;

QY 77 PCIPNCHNGTCEIS-EAYRGDTFIGYCKPRGFNGHCOHNINECEAECCRNIGICT 135
 Db 1 PCSPNPCYNDACVLTLDIQRGDIFREYICQCPVGSIGHCETETNYNLD--GEYMF 57
 QY 136 DLVANYSCCEPGEF--MGRNCQKCSGHLGIEGIIISNOQITASSNHRALFGLQKWPY 193
 Db 58 TAVPNTAVPTPAVTPDLNSNLASRCSQTQGMGGAIDQSIASSVYMGFLQRMGP 117
 QY 194 ARLNKGLINATTAENDMPWIOINLQKRAVYGVITOGAKRISPEYIKSYKTAYSND 253
 Db 118 ARLYRGIVNATVASYDSKPMIYQVNLKRMVSGVMTQASRAGRAREYLTFFKVAVSLD 177
 QY 254 GK 255
 Db 178 GR 179

RESULT 15

061240 PRELIMINARY; PRT; 2352 AA.

AC 061240; 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE HRNOTCH PROTEIN.
 GN HRNOTCH.

OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyrosidae; Halocynthia.
 NCBI_TaxID=7729;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Hori S., Saitoh T., Matsumoto M., Makabe K.W., Nishida H.;
 RL Dev. Genes Evol. 207:371-380(1997).
 DR EMBL; AB001327; BAA2571.1; -;
 DR HSSP; P00740; IEDM.
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000742; -;
 DR INTERPRO; IPR000800; -;
 DR INTERPRO; IPR001438; -;
 DR INTERPRO; IPR001881; -;
 DR INTERPRO; IPR002110; -;
 DR PFAM; PF00008; EGF_32.
 DR PFAM; PF00023; ank; 6.
 DR PFAM; PF00066; notch; 3.
 DR PRINTS; PR00010; EGRLOD.
 DR PROSITE; PS00010; ASX_HYDROXYL; 18.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_28.
 DR PROSITE; PS01186; EGF_2; 22.
 DR PROSITE; PS01187; EGF_CA; 18.
 KW Glycoprotein; EGF-like domain.
 SQ SEQUENCE 2352 AA; 252622 MW; 13DB1C056BBD08D CRC64;

Query Match

Best Local Similarity 12.8%; Score 341; DB 5; Length 2352;

Matches 61; Conservative 17; Mismatches 45; Indels 22; Gaps 4;

QY 26 CNPNCENGGICLGLADDSFSCCEPGEFAGPNCSSVVEVASDEKPTSAEPCIPNCHN 85
 Db 720 CATDPCQNGGICTSGI--NSYNCACPAKYTGVCN-----TELSPCVNPNCEN 765
 QY 86 GGTCEISEAYRGDTFIGYCKPRGFNGHCOHNINECEAEPCRNIGICTDLVANYSC 145
 Db 766 GATCOESADY-----LAVYCCPGEFGRPGTCATDINECVNSPCXNGGCTNLVPGYCTC 820
 QY 146 PGEFGRNCQY--KCSGHLGIEGG 167
 Db 821 SQGFTGKDCDIDDCSSNPCLNGG 845

Search completed: May 23, 2001, 06:20:42
 Job time: 436 sec


```

FT      /note="discoidin I/factor VIII-like domain 1"
FT      319..476
FT      /label= Discoidin-2
FT      /note="discoidin I/factor VIII-like domain 2"
XX      W06640769-A1.
XX      19-DEC-1996.
XX      05-JUN-1996; 96MO-US09456.
XX      07-JUN-1995; 95US-0480229.
XX      (PROG-) PROGENITOR INC.
XX      (UYVA-) UNIV VANDERBILT.
XX      Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX      WPI, 1997-052233/05.
XX      N-PSDB; T47338.
XX      New developmentally regulated endothelial cell locus-1 (del-1) gene
XX      - used to develop prods. for the diagnosis and treatment of cancer
XX      and conditions involving abnormal angiogenesis
XX      Claim 3; Fig 6; 137pp; English.
XX      Murine Del-1 (W10364) is the polypeptide product of the murine
XX      developmentally-regulated endothelial cell locus-1 (del-1) gene
XX      (T47338). It shows 94% amino acid homology to the human Del-1
XX      protein (W10365). Structurally, members of this novel gene family
XX      contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
XX      domains. Del-1 is expressed in endothelial and certain tumor
XX      cells. Its ability to inhibit vascular formation allows its used
XX      as an anti-angiogenic agent. It can be used as a tumor marker.
XX      to identify del-1 binding partners, and to modulate endothelial
XX      cell growth and blood vessel formation. Recombinant Del-1 can be
XX      produced in transformed host cells utilizing vectors incorporating
XX      del-1 nucleic acids.
XX      Sequence 480 AA:
XX
Query Match      100.0%; Score 2669; DB 18; Length 480;
Best Local Similarity 100.0%; Pred. No. 2.8e-148;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 MKHLVAMLLVGLSLGVPQFGKGDICNPNCENGICISGLADSFSCPCPGFAGPNC 60
DB      1 mkhlvaawllvglslygvpqfggdicnppcengicisgladsfscpcpgfagpncs 60
OY      61 SYVEVASDEEKPISAGPCIPNCPCHNGTCEISEAYRGDTFTGYVCKCPRGFNGICQHNI 120
DB      61 syvevasdeekpitsagpcipncpchngtceiseayrgdtftigyvckcprgfgngicqnhl 120
OY      121 NCCBAPCNGGICIDIVANVSCPCRGEMGNCOYKCSGHGIGTISNOQTASSNH 180
DB      121 nccbapcnggicidivanyvscpcrgemgncoykcsghgigtisnoqtassnh 180
OY      121 nccbapcnggicidivanyvscpcrgemgncoykcsghgigtisnoqtassnh 180
DB      121 nccbapcnggicidivanyvscpcrgemgncoykcsghgigtisnoqtassnh 180
OY      181 RALFGLOKWPYPYARLNKGLINAWTAENDRPWIOINLRKMBVTGVTIOGAKRISP 240
DB      181 ralfglokwpypyarlnkglinawtaendwpwiqinlqrkmvtgvtiogakrissp 240
OY      241 ETKSTKLAISMDGKTWAMYKVKGTNEEMVFRGNVNTPYANSEFPPIKAQYRLYPOI 300
DB      241 etkstklaismdgktwamykvkgtneemvfrgnvntpyansefppikaqyryllypqi 300
OY      301 CRRHCTLRHELIGCELSGCEPGLMKSGHIDYOTTASSVFRITLNMDFTEPRKARLDK 360
DB      301 crrhctlrheligcelsgcseplgmksghidyoTTASSVFRITLNMDFTEPRKARLDK 360
OY      301 crrhctlrheligcelsgcseplgmksghidyoTTASSVFRITLNMDFTEPRKARLDK 360
DB      301 crrhctlrheligcelsgcseplgmksghidyoTTASSVFRITLNMDFTEPRKARLDK 360
OY      361 GGRVNAWTSIGNHDSQWLOVDLVPYTKVTGITOGAKDFGVQFVSGFKLAYSNDGEHMM 420
DB      361 ggrvnaawtsignhdsqwlqvdlvpytkvtglttgakdfghvqfvgysklyasndgehmm 420

```

```

DB      361 ggrvnaawtsignhdsqwlqvdlvpytkvtglttgakdfghvqfvgysklyasndgehmm 420
OY      421 VHODEKORCKDYQGNFNDTHRKKNVIDPPIYARFIRILPMSWGRITLRSELGCAEE 480
DB      421 vhddekqrckdyqgnfndthrkknvidppiyarfirilpmswgrilrseellgcaeee 480
XX      RESULT 2
XX      W94683
XX      ID W94683 standard; Protein; 480 AA.
XX      AC W94683;
XX      DT 05-MAY-1999 (first entry)
XX      DE Murine Del-1 protein.
XX      KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
XX      KW discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
XX      KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
XX      KW angiogenesis.
XX      OS Mus sp.
XX      PN US5877281-A.
XX      PD 02-MAR-1999.
XX      PE 05-JUN-1996; 96US-0659235.
XX      PR 05-JUN-1996; 96US-0659235.
XX      PR 07-JUN-1995; 95US-0480229.
XX      PA (PROG-) PROGENITOR INC.
XX      PA (UYVA-) UNIV VANDERBILT.
XX      PI Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX      WPI, 1999-189720/16.
XX      DR N-PSDB; X18506.
XX      FT Del-1 polypeptide sequences - useful for treatment of cancer,
XX      PS diabetic retinopathy, rheumatoid arthritis and endometriosis
XX      PS Claim 1; Fig 3; 73pp; English.
XX
CC      The present sequence is murine developmentally-regulated endothelial
CC      cell locus 1 (del-1). The protein has epidermal growth factor (EGF) like
CC      domains and discoidin I/factor VIII-like domains. The Del-1 proteins
CC      have an inhibitory effect on angiogenesis (blood vessel growth), this
CC      activity may be useful clinically to prevent neovascularisation of
CC      tissues such as tumor nodules and prevention of metastases. The anti-
CC      angiogenic activity of Del-1 may be used to treat abnormal conditions
CC      that result from angiogenesis, including cancer, diabetic retinopathy,
CC      rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC      angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC      stroke, wound healing and peripheral vascular disease. Del-1 is also
CC      useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC      integrin, and is an apoptosis inducer.
XX      Sequence 480 AA:
XX
Query Match      100.0%; Score 2669; DB 20; Length 480;
Best Local Similarity 100.0%; Pred. No. 2.8e-148;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 MKHLVAMLLVGLSLGVPQFGKGDICNPNCENGICISGLADSFSCPCPGFAGPNC 60
DB      1 mkhlvaawllvglslygvpqfggdicnppcengicisgladsfscpcpgfagpncs 60
OY      61 SYVEVASDEEKPISAGPCIPNCPCHNGTCEISEAYRGDTFTGYVCKCPRGFNGICQHNI 120
DB      61 syvevasdeekpitsagpcipncpchngtceiseayrgdtftigyvckcprgfgngicqnhl 120

```

Db 61 svveasdeekpksapciipnchngtclseayrgdtfigyvcckprgfnghcqhni 120
 QY 121 NECEAPCRNGICTDILVANSCECPGEFMGRNCQKCSGHLIEGGIISNOQITASSNH 180
 Db 121 neceapcrngictdilanyscecpgefmgrncqkcsghlgegglisnqglitassnh 180
 QY 181 RALFGLQKWPYVARLNKGLINAWTAENDRWPVQINLQKRMVTVITOGAKRIGSP 240
 Db 181 ralfglqkwpvyarlknkglinaawtaendrwplqinlqkrmvrtvltgqakrigrsp 240
 QY 241 EYIKSKRTAIVSNDGKTWAMKVKGTNEEMFRGNVDNNPYNASTFPPIKAQYVRLYQI 300
 Db 241 eyskrtayvsnkgktwamkvkgtneemfrgnvdnnpynastfprikaqyvrlyqi 300
 QY 301 CRRHCTLMELLGCELSGCESEPLGMSGHIQDYQITASSVFRTLLNMDFTWEPKRALDK 360
 Db 301 crrhctlmellgcelsgcesepigmksgihdqyqitassvfrtllnmdftwepkraldk 360
 QY 361 QGKVAWMTSGHNDOSQWLDVLPVTKVGTITOGAKDFGHQVFGSYKLAYSNDGEHM 420
 Db 361 qgkvnawtsgnhndosqlvdlvpktvgtltcgakdfghvqfvsyklaysndgehm 420
 QY 421 VNODEKORKDKVFOGNFNDTHRKKNVIDPPIYARFIRILPMSWYGRITLRSELLGCAEE 480
 Db 421 vnodekorkdkvfgnfnndthrkknvidppiyarfilrlpmswygritlrseellgcaeee 480
 RESULT 3
 W94685 ID W94685 standard; Protein: 513 AA.
 AC W94685;
 XX 05-MAY-1999 (first entry)
 DE Human Del-1 protein.
 KW Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KW discoidin 1; factor VIII-like domain; epidermal growth factor; EGF;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW angiogenesis.
 OS Homo sapiens.
 XX
 PN US5877281-A.
 XX 02-MAR-1999.
 PD
 XX 05-JUN-1996; 96US-0659235.
 PR 05-JUN-1996; 96US-0659235.
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Hogan B, Quettermous T, Snodgrass HR, Zupancic TJ;
 XX WPI. 1999-189720/16.
 DR N-PSDB; X18508.
 XX
 PT Del-1 polypeptide sequences - useful for treatment of cancer,
 XX diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Claim 5; Fig 4; 73pp; English.
 CC The present sequence is human developmentally-regulated endothelial
 CC cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like
 CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
 CC have an inhibitory effect on angiogenesis (blood vessel growth), this
 CC activity may be useful clinically to prevent neovascularisation of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions

CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC integrin, and is an apoptosis inducer.
 XX
 SQ Sequence 513 AA;
 Query Match 94.4%; Score 2519.5; DB 20; Length 513;
 Best Local Similarity 94.2%; Pred. No. 1,5e-139;
 Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;
 QY 1 MKHLVAMWLVLSGVPOFGKGDICNPENCGICLSGLADSFSECEPGEFAPNGS 60
 Db 33 mkrsavwvlvlgslgvpoqfgkdicdpnpenngiclpjglavgsfscpepdgftdncs 92
 QY 61 SVVEVASDEEKPTSGAPCIPNPNCHNGTCEISEAYRGDTFIGYVCCKPRFNGIHQCHNI 120
 Db 93 svvevasdeekptsgapcnpnchngtceiseayrgdtfigyvcckprgfnghcqhni 152
 QY 121 NECEAPCRNGICTDILVANSCECPGEFMGRNCQKCSGHLIEGGIISNOQITASSNH 180
 Db 153 neceapcrngictdilanyscecpgefmgrncqkcsghlgegglisnqglitassnh 212
 QY 181 RALFGLQKWPYVARLNKGLINAWTAENDRWP-VQINLQKRMVTVITOGAKRIGS 239
 Db 213 ralfglqkwpvyarlknkglinaawtaendrwvqinlqkrmvrtvltgqakrigrs 272
 QY 240 PEYIKSKRTAIVSNDGKTWAMKVKGTNEEMFRGNVDNNPYNASTFPPIKAQYVRLYQI 299
 Db 273 peyikrtayvsnkgktwamkvkgtneemfrgnvdnnpynastfprikaqyvrlyqi 332
 QY 300 ICRHCTLMELLGCELSGCESEPLGMSGHIQDYQITASSVFRTLLNMDFTWEPKRALD 359
 Db 333 icrrhctlmellgcelsgcesepigmksgihdqyqitassvfrtllnmdftwepkrald 392
 QY 360 QGKVAWMTSGHNDOSQWLDVLPVTKVGTITOGAKDFGHQVFGSYKLAYSNDGEHM 419
 Db 393 qgkvnawtsgnhndosqlvdlvpktvgtltcgakdfghvqfvsyklaysndgehm 452
 QY 420 MNODEKORKDKVFOGNFNDTHRKKNVIDPPIYARFIRILPMSWYGRITLRSELLGCAEE 479
 Db 453 mnodekorkdkvfgnfnndthrkknvidppiyarfilrlpmswygritlrseellgcaee 512
 QY 480 E 480
 Db 513 e 513
 RESULT 4
 W10365 ID W10365 standard; Protein: 481 AA.
 XX
 AC W10365;
 XX 03-MAY-1997 (first entry)
 DE Human developmentally-regulated endothelial cell locus-1 protein.
 XX
 KW Del-1; developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnerability; bone formation; diagnosis; therapy.
 OS Homo sapiens.
 XX
 XX key Location/Qualifiers
 FH Peptide 3..23
 FT /label= Sig_peptide
 FT Protein 24..481

```

FT      Domain      /label= Mat.-protein
FT      26..59
FT      /label= EGF-1
FT      /note= "epidermal growth factor-like domain 1"
FT      78..116
FT      Domain      /label= EGF-2
FT      /note= "epidermal growth factor-like domain 2"
FT      123..154
FT      /label= EGF-3
FT      /note= "epidermal growth factor-like domain 3"
FT      158..315
FT      Domain      /label= Discolidin-1
FT      /note= "discolidin I/factor VIII-like domain 1"
FT      320..477
FT      Domain      /label= Discolidin-2
FT      /note= "discolidin I/factor VIII-like domain 2"
FT      480..540
FT      19-DEC-1996.
FT      W09640769-A1.
FT      19-DEC-1996.
FT      05-JUN-1996; 96MO-US09456.
FT      07-JUN-1995; 95US-0480229.
FT      (PROG-) PROGENITOR INC.
FT      (UYVA-) UNIV VANDERBILT.
FT      Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
FT      WPI; 1997-052233/05.
FT      N-PSDB; T47343.
FT      New developmentally regulated endothelial cell locus-1 (del-1) gene
FT      - used to develop prods. for the diagnosis and treatment of cancer
FT      and conditions involving abnormal angiogenesis
FT      Claim 4; Fig 6; 137pp; English.
FT      Human Del-1 (W10365) is the polypeptide product of the human
FT      developmentally-regulated endothelial cell locus-1 (del-1) gene
FT      (747343). It shows 94% amino acid homology to the mouse Del-1
FT      protein (W10364). Structurally, members of this novel gene family
FT      contain 3 EGF-like domains and 2 discolidin I/factor VIII-like
FT      domains. Del-1 is expressed in endothelial and certain tumour
FT      cells. Its ability to inhibit vascular formation allows its used
FT      as an anti-angiogenic agent. It can be used as a tumour marker,
FT      to identify Del-1 binding partners, and to modulate endothelial
FT      cell growth and blood vessel formation. Recombinant Del-1 can be
FT      produced in transformed host cells utilizing vectors incorporating
FT      del-1 nucleic acids.
FT      Sequence 481 AA;
SQ
Query Match          94.1%; Score 2512.5; DB 18; Length 481;
Best Local Similarity 94.0%; Pred. No. 3,6e-139;
Matches 457; Conservative 9; Mismatches 19; Indels 1; Gaps 1;
QY      1 MKHLVAMLVGLSLGVPOFGKGDICNPNCENGICISGLHDSFSCPECFAGNCS 60
DB      1 mksvavwllvgislgvpfgkygdicdpncengiclpjlagvisfscpecdgfdpncs 60
QY      61 SYVEVASDEKPTISAGPCIPNCHNGTGEISEAVRGOTFFIGYCKPCRFNGIHCOHNT 120
DB      61 syvevasdeepsagpcipnchngtgeiseavrgotffigyckprfngihcqhnl 120
QY      121 NECEAECRMGICITDLVANYSCCEGPEFMGNCOYKCSGHIGIEGGIISMOQTASSNH 180
DB      121 neceaeckngicidlvanysecepgelmgncqkcsqplglegdlisnqqltaasth 180
QY      181 RALFGLQKWVYARLNKKGLINAMTAENDRW-FWIOINLQKRWYTGVTQCAKRIGS 239
DB      181 ralfglqkwyppyarlnkkglinaawtaendrwvriqinlqkrmcvtgvtlgakrigrs 240

```

```

DB      181 ralfglqkwyppyarlnkkglinaawtaendrwvriqinlqkrmcvtgvtlgakrigrs 240
QY      240 PEYIKSYKRIAYSNDKQTMAMVKVGTNEEMVFRGNVDNPTPYANSFTPPKAKQVRLYPO 299
DB      241 peyiksykriaysndkqtmamkvgtneemvfrgnvdnptpyansftppkakyvrlypo 300
QY      300 ICRHICTIRMLLCCESGSEPLGKMSGHIDYOTITASSVVRTINMDMFTWEPKALD 359
DB      301 icrhictirmllycelsgscseplgmsghidytitassvrrtinmdmftwepkarld 360
QY      360 KQKVNAMTSGHNDQSOMLQVDLVPFKVGLITOGAKDPGHVOFVGSYKLIAYSNDEBHW 419
DB      361 kgkvnamtsgnhndqsgmlqvdlvpfkvglltogaakdfghvqfvgysklyaysndeghw 420
QY      420 MYHDEKORAKDVQGNFNDTHRKAVIDPPYIARTIRLPSMTGRITLNSLLCAEE 479
DB      421 myhdekorkdvqgnfnndthrkavidppiyartrilpswtgritlslsllgcee 480
QY      480 E 480
DB      481 e 481
RESULT 5
W94697
ID      W94697 standard; Protein; 321 AA.
XX      W94697;
AC      05-MAY-1999 (first entry)
DT      Human milk fat globule protein MFG-E8.
DE      XX
KW      Del-1; developmentally-regulated endothelial cell locus 1; cancer;
KW      discolidin I; factor VIII-like domain; epidermal growth factor; EGF;
KW      diabetic retinopathy; rheumatoid arthritis; endometriosiis;
KW      angiogenesis.
OS      Homo sapiens.
XX      OS
XX      PN      US5877281-A.
XX      PD      02-MAR-1999.
XX      PF      05-JUN-1996; 96US-0659235.
XX      PR      05-JUN-1996; 96US-0659235.
XX      PR      07-JUN-1995; 95US-0480229.
XX      PA      (PROG-) PROGENITOR INC.
XX      PA      (UYVA-) UNIV VANDERBILT.
XX      PI      Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
XX      DR      WPI; 1999-169720/16.
XX      PT      Del-1 polypeptide sequences - useful for treatment of cancer,
XX      diabetic retinopathy, rheumatoid arthritis and endometriosiis
XX      PS      Example; Column 63-64; 73pp; English.
CC      The present sequence represents human milk fat globule protein MFG-E8,
CC      which has homology to the developmentally-regulated endothelial cell
CC      locus 1 (Del-1). The Del-1 protein has epidermal growth factor (EGF) like
CC      domains and discolidin I/factor VIII-like domains. The Del-1 proteins
CC      have an inhibitory effect on angiogenesis (blood vessel growth), this
CC      activity may be useful clinically to prevent neovascularisation of
CC      tissues such as tumour nodules and prevention of metastases. The anti-
CC      angiogenic activity of Del-1 may be used to treat abnormal conditions
CC      that result from angiogenesis, including cancer, diabetic retinopathy,
CC      rheumatoid arthritis and endometriosiis. Since Del-1 promotes
CC      angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC      stroke, wound healing and peripheral vascular disease. Del-1 is also

```



```

XX 05-JUN-1996; 9605-0659235.
PF
XX
PR 05-JUN-1996; 9605-0659235.
PR 07-JUN-1995; 9505-0480229.
XX
PA (PROG-) PROGENITOR INC.
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Quercernous T, Snodgrass HR, Zupancic TJ;
DR WPI, 1999-189720/16.
DR N-PSDB; X18507.
XX
PT Del-1 polypeptide sequences - useful for treatment of cancer,
PT diabetic retinopathy, rheumatoid arthritis and endometriosis
XX
PS Claim 3; Column 71-72; 73pp; English.
XX
CC The present sequence is truncated murine developmentally-regulated
CC endothelial cell locus 1 (Del-1). The protein has epidermal growth factor
CC (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1
CC proteins have an inhibitory effect on angiogenesis (blood vessel growth),
CC this activity may be useful clinically to prevent neovascularisation of
CC tissues such as tumour nodules and prevention of metastases. The anti-
CC angiogenic activity of Del-1 may be used to treat abnormal conditions
CC that result from angiogenesis, including cancer, diabetic retinopathy,
CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
CC stroke, wound healing and peripheral vascular disease. Del-1 is also
CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
CC integrin, and is an apoptosis inducer.
XX
SQ Sequence 221 AA;

```

Query Match 46.5%; Score 1242; DB 20; Length 221;
 Best Local Similarity 98.2%; Pred. No. 1,66-65;
 Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MKHLVAMLLVGLSLGVPQFGKGDICNPNCENGICLSGLADSPSCPEGEFAGPNC60
Db 1 mkhlvaawllvglslygpyqgkqgdiicnpncengiclsjladstscpeefagpnc60
QY 61 SYVEVASDEEKPTSGAPCIIPNCPHNGTCEISEAYRGDFIYGVCCKPRGFNGIHCOHNI 120
Db 61 syvevasdeekptsagpcipncphngtceiseayrgdfiygvcckprgfnghcqhni 120
QY 121 NECEAEPCRRNGICTDLYVANYSCCEPGEFGRNCQYKCSGHLIEGGIISNOQITASSNH 180
Db 121 neceaepecrngictdlyvanyscepgefgrncqykcsghlgegglisnqitassnh 180
QY 181 RALFGLQKWYPYARLNKKGLINAMTAENDRPMWIQINL 220
Db 181 ralfglqkwypyaalnkkglinawtaendrpmwiyqivcr 220

```

RESULT 8
 Y94454
 ID Y94454 standard; Protein; 426 AA.
 XX
 AC Y94454;
 XX
 DT 11-SEP-2000 (first entry)
 XX
 DE Mouse lactadherin protein.
 XX
 KW Human; lactadherin; MGF-E8; anti-tumour; immune response;
 KW exosome; dendritic cell.
 OS Mus sp.
 XX
 FH Key Location/Qualifiers

```

FT Peptide 1..22
FT /label= Secretion_signal
FT Protein 23..426
FT /label= Lactadherin
FT Binding-site 87..89
FT /label= Integrin_binding_site
FT Misc-difference 93..111
FT /note= "encoded by Gn"
XX
XX EPI004664-A1.
XX
XX 31-MAY-2000.
XX
XX 24-NOV-1998; 98EP-0402925.
XX
XX 24-NOV-1998; 98EP-0402925.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX (CUR1-) INST CURIE.
PA
DR WPI: 2000-352597/31.
DR N-PSDB; A27141.
XX
XX Chimeric isolated (human) lactadherin polypeptide that functions as an
XX adaptor of cross-priming to eliminate pathogenic antigens, e.g. in
XX tumors -
XX
XX Disclosure; Page 12; 20pp; English.
XX
XX Lactadherin protein was found in exosomes produced by dendritic cells.
XX The protein is involved in the phagocytosis of particulate antigens by
XX dendritic cells. Exosomes produced by dendritic cells exposed to
XX tumour antigens induce potent immune responses. Lactadherin or variants
XX of it may be used in the mediation of an immune response. Variants of
XX lactadherin may be used for inhibition and/or stimulation of the
XX cross-priming of antigens and stimulation of the phagocytosis of
XX antigens by dendritic cells. Compositions derived from lactadherin can
XX also be used to monitor an immune response, more specifically a CTL
XX (cytotoxic T-lymphocyte) response and also to produce CTLs specific
XX for a selected antigen. The present sequence is the mouse
XX lactadherin protein.
XX
SQ Sequence 426 AA;

```

Query Match 43.0%; Score 1149; DB 21; Length 426;
 Best Local Similarity 47.6%; Pred. No. 7,66-60;
 Matches 217; Conservative 70; Mismatches 113; Indels 56; Gaps 7;

```

QY 23 GDICNPNCENGICLSGLADSPSCPEGEFAGPNCSSVVEVASDEEKPTSGAPCIIPNP 82
Db 23 gdicnpncengiclsjladstscpegefagpncssvvevasdeekptsagpcipnp 82
QY 25 gqfcdssalngyctcltg-qdndlyclcpegftgvcne-----lergpcspnp 72
Db 25 gqfcdssalngyctcltg-qdndlyclcpegftgvcne-----lergpcspnp 72
QY 83 CHNGSTCEIS-EAYRGDFIYGVCCKPRGFNGIHCOHINECEAPCNGICITDLYVANY 141
Db 83 chngstceis-eayrgdfiygvcckprgfnghcqhineceapcngicitdlyvany 141
QY 73 cynaakclvltidtgqdfiteyloqcpvgsgh----- 106
Db 73 cynaakclvltidtgqdfiteyloqcpvgsgh----- 106
QY 142 SCECPGEFGRNCQYKCSGHLIEGGIISNOQITASSNHRALFGLQKWYPYARLNKKGL 201
Db 142 scecpgefgrncqykcsghlgegglisnoqitassnhrralfglqkwypyarlnkkgl 201
QY 107 -----cetecstqjmeggaladqiasyyymfmgldqrgvpeparilytql 154
Db 107 -----cetecstqjmeggaladqiasyyymfmgldqrgvpeparilytql 154
QY 202 INAMTAENDRPMWIQINLQKRMRYTVITOGAKRIGSPKSYKIAVSDGKTWAMYK 261
Db 202 inamtaendrpmwiqinlqkrmrvtvitogakrigspkisykiavsdgktwamyk 261
QY 155 vnavhasnydskpviqvnllrkmyrsgymtgasragraeylktfkwaysldgkffeliq 214
Db 155 vnavhasnydskpviqvnllrkmyrsgymtgasragraeylktfkwaysldgkffeliq 214
QY 262 -VKGTNEEMFVRGNVDNTPFANSTFPIKQYRLYQICRRICITLMEITGELSGCS 320
Db 262 -vkgtneemfvrngvndntpfanstfpiikqyrllyqicrricitlmetelsgscs 320
QY 215 desgdkc--flgndnslkvnmfnplcleagylrlyvpschrgctrlfelglcelhgc 272
Db 215 desgdkc--flgndnslkvnmfnplcleagylrlyvpschrgctrlfelglcelhgc 272
QY 321 EPLGKRSGHIDYQITASSVFRTLNMDFTWEPKRLDKOGKVNAMTSGHNDOSOMLOY 380
Db 321 eplgkrsghidoyitassvfrtlnmdftwepkrlldkogkvnamtsghndosomloy 380
QY 273 eplgikntltpdsqmsasssykltwnlrafwyphlgrldnggklnawtaagsnsakewlqv 332
Db 273 eplgikntltpdsqmsasssykltwnlrafwyphlgrldnggklnawtaagsnsakewlqv 332

```

```

OY 381 DLVPTKVYGIITOGAKDFGHVQFVGSYKLAYSNDGEHMVYHODEKORKKRVFGNEDND 440
  || :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 333 dltgrgrytgitltggarfgnhtqyvesykvahsddvgwtvy--eegsskxvfqndnn 390
OY 441 THKKNVIDPPIYARFIRILPMSWGRITLRSRLGCG 476
  :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 391 shknlfekpmaryrvlpswvhnrlrlfllg 426

RESULT 9
R77252
ID R77252 standard; Protein: 387 AA.
XX
AC R77252;
XX
DT 21-NOV-1995 (first entry)
XX
DE HMFg 46 kDa antigen.
XX
KM HMFg; human milk fat globule; antigen; immunogen; vaccine; virucide;
XX epithelium; tumor; breast cancer; monoclonal antibody; Mab.
XX
OS Homo sapiens.
XX
PN W09515171-A.
XX
PD 08-JUN-1995.
XX
PE 05-DEC-1994; 94WO-US13967.
XX
PR 03-DEC-1993; 93US-0162402.
XX
PA (CANC-) CANCER RES FUND CONTRA COSTA.
XX
PI Certani RL, Larocca DJ, Peterson JA;
XX
DR WPI: 1995-215151/28.
XX N-PSDB; 091198.
XX
DR 46 kD apparent molecular weight human milk fat globule antigen -
XX used in assays to determine the presence of a cancerous tumour of
XX epithelial origin, and in a vaccine against neoplastic tumours
XX
PS Claim 6; Page 46-47; 68pp: English.
XX
CC A complete cDNA sequence for the 46 kDa HMFg antigen, a major
XX component of the apical surface of the normal breast epithelial
XX cell, was obtained by PCR and RACE methods. cDNA clones can be
XX used to prepare MAb for use in immunotherapy, immunohistopathology,
XX prognosis, diagnosis, imaging and therapy. Recombinant antigen can
XX be expressed in prokaryotic or (glycosylated) in eukaryotic cells.
XX
SQ Sequence 387 AA;

Query Match 38.3%; Score 1023; DB 16; Length 387;
Best Local Similarity 48.9%; Pred. No. 1.5e-52;
Matches 196; Conservative 55; Mismatches 108; Indels 42; Gaps 4;
OY 78 CIPNCHNGTTC-EISEAYRGDTFTGYVCKPRGFNGIHCOHINIECBAPCRNGGICTD 136
  | :|||||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 27 cskpchnhgltceisqevrvgdfpsylctclkgaygnh----- 65
OY 137 LVANVSCCPGPEFMGRNQYKSCSHLGTGEGITISNOOITASSNHRALFGLQKWYYARL 196
  | :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 66 -----celkceveplgmengniansqiaassvvtflgldhwpelarl 108
OY 197 NKGGLINMTAENDRMFMIQINLORKRMVGVITOGAKRIGSPETKSYKVIASNDGKT 256
  | :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 109 nragvnaewtpsdnddpwlgvnlrlrmvlgvvtgagrsilasheykafkavayslnghe 168
OY 257 W-AMTKVKGITNEBVFRRGVNDNNTFYANSFTPIKAQVVRKLYPQICRKHCTLRMELLGCE 315
  : :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|

```

```

Db 169 fdfldhvnkknkhekv--gnwnknavhvnlfetpveaqvrvlyptschactllrfellgce 226
OY 316 LSGCSEPLGKSGKHIDYQYTASSVFRTLMKDMFTWEPKARLDDKQKVAWTSGHNDOS 375
  | :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 227 lngcanplgiknnsipdkqitasssyktwglhlfswpsyarldkqgnfnawagsygn 286
OY 376 QMLOVDLVPYTKVGIITOGAKDFGHVQFVGSYKLAYSNDGEHMVYHODEKORKKRVFG 435
  | :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 287 gwlqyvdlygsketvgtltggarfnfsgvqfvasykvaynsdanasvteyqdpdrtgsklfpg 346
OY 436 NFDNDTHKKNVIDPPIYARFIRILPMSWGRITLRSRLGCG 476
  | :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db 347 nwdhshknlfecplaryrvrlpswvhnrlrlfllg 387

RESULT 10
Y94453
ID Y94453 standard; Protein: 387 AA.
XX
AC Y94453;
XX
DT 11-SEP-2000 (first entry)
XX
DE Human lactadherin protein.
XX
KM Human; lactadherin; MGF-E8; anti-tumour; immune response;
XX exosome; dendritic cell.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Protein /label= Secretion_signal 24..387
FT /label= Lactadherin 46..48
FT Binding-site /label= Integrin_binding_site
XX
PN EP1004664-A1.
XX
PD 31-MAY-2000.
XX
PF 24-NOV-1998; 98EP-0402925.
XX
PR 24-NOV-1998; 98EP-0402925.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX (CUR-) INST CURIE.
XX
DR WPI: 2000-352597/31.
XX N-PSDB; A27140.
XX
PT Chimeric isolated (human) lactadherin polypeptide that functions as an
PT adaptor of cross-priming to eliminate pathogenic antigens, e.g. in
PT tumors -
XX
XX Example 3; Page 12; 20pp: English.
XX
CC Lactadherin protein was found in exosomes produced by dendritic cells.
XX The protein is involved in the phagocytosis of particulate antigens by
XX dendritic cells. Exosomes produced by dendritic cells exposed to
XX tumour antigens induce potent immune responses. Lactadherin or variants
XX of it may be used in the mediation of an immune response. Variants of
XX lactadherin may be used for inhibition and/or stimulation of the
XX cross-priming of antigens and stimulation of the phagocytosis of
XX antigens by dendritic cells. Compositions derived from lactadherin can
XX also be used to monitor an immune response, more specifically a CTL
XX (cytotoxic T-lymphocyte) response and also to produce CTLs specific
XX for a selected antigen. The present sequence is the human
XX lactadherin protein.
XX
SQ Sequence 387 AA;

```

Query Match:	38.3%:	Score 1023:	DB 21:	Length 387:
Best Local Similarity	48.9%:	Pred. No. 1.5e-52:		
Matches 196:	Conservative 55:	Mismatches 108:	Indels 42:	Gaps 4:
Qy 78	CIPNPGNGTGC-EISEAYRGDTFYGVCPCRGFNGIHCOHNTNECAEPCRRNGICTD	136		
Db 27	csknpcnngljceelsqevrgdvfpysctctclkyagph-----	65		
Qy 137	LVANYSECPGGERMGNCGKCSGHLEGGIISNQOTTASSNHRALFGLQKWYPYARL	196		
Db 66	-----cetkeveplmgengniansqdaasvvtltlqghwvpealrt	108		
Qy 197	NKGLINAMTAENDRPWIIQINLQKMRVTCVITOGAKRIGSEPYIKSYIAVSDGKT	256		
Db 109	ntagwvnaawpspsnddpwilyvlltrmvltgvyvgasstrlasheytkafkayvsinghe	168		
Qy 257	W-AAYKKKGNENMVRGNVDNNTPYANSFTRPIKACQVRLRPDICHRRHDTLMEIIGE	315		
Db 169	fdffhvdnrxkhefv--gnmknahvnhlfepveaqyvrlypschactactlrfellge	226		
Qy 316	LSGSEPELKKSGCHIDDOYQTASSVPRITLNDMFTRERKARLDKQKVAAMTSGHNDOS	375		
Db 227	Inganapliiknsipdkqkllasssyktwqlhltswmpsyarldkqgnflawagasynd	286		
Qy 376	QWLOVDLVETKVTGIIITOGAKDFGVQVSGYKLAASNDSEHMVHODEKORKDKVFOG	435		
Db 287	gwldvqdsakevrgylltgarnfgsvglfvaasykvaandsanvteyqdpbrtgsklfpg	346		
Qy 436	NFQNDPRKRVNDIPRYARFRIILPMSWYGRITLRSSELLGC	476		
Db 347	nwdhskhknlfetcpilaryvllpwawmtairtellegc	387		
RESULT 11				
W98891	W98891 standard; peptide: 320 AA.			
XX AC	W98891:			
XX DT	05-MAY-1999 (first entry)			
XX DE	Peptide SEQ ID NO:20 from US5877281.			
XX OS	Unidentified.			
XX PN	US5877281-A.			
XX PD	02-MAR-1999.			
XX PF	05-JUN-1996; 960S-0659235.			
XX PR	05-JUN-1996; 960S-0659235.			
XX XX	07-JUN-1995; 95US-0480229.			
XX PA	(PROG-) PROGENITOR INC.			
XX PI	(UTVA-) UNIV VANDERBILT.			
XX PI	Hogan B, Quertermous T, Snodgrass HR, Zupancic RJ;			
XX DR	WPI; 1999-189720/16.			
XX PT	Del-1 polypeptide sequences - useful for treatment of cancer,			
XX PT	diabetic retinopathy, rheumatoid arthritis and endometriosis			
XX PS	disclosure; Column 61-62; 73pp: English.			
XX CC	The present invention describes developmentally-regulated endothelial			

[illegible]

FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site /note= "potential N-glycosylation site"
FT Cleavage-site /note= "thrombin cleavage site"
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site /note= "potential N-glycosylation site"

MO9630546-A1.

03-OCT-1996.

22-MAR-1996: 96WO-US03881.

24-MAR-1995: 95US-0410488.

(SCRI) SCRIPPS RES INST.

Gandhille S, Greengard J, Griffin JH;

WPI, 1996-455389/45.

N-PSDB: T33942.

Detection of Factor V gene mutation - by PCR amplification to identify exon 10 guanine 205 or 1691 to adenine substitution, which results in activated Protein C resistance

Disclosure: Page 112-127; 175pp; English.

The amino acid sequence (W04254) of human Factor V was deduced from a cDNA clone (T33942). Factor V is involved in the blood coagulation process. Activated protein C (APC) resistance, characterised as a poor anticoagulant response to APC in plasma, is a result of a guanine to adenine substn. at position 1691 of the Factor V cDNA.

Sequence 2224 AA;

Query Match 25.8%; Score 689; DB 17; Length 2224;

Best Local Similarity 42.7%; Pred. No. 2,2e-32; Matches 144; Conservative 55; Mismatches 122; Indels 16; Gaps 4;

144 ECPGEFMRNCOYKCSGHLGIEGIIISNOQTASSNHRALFGLQKWPYARLNKKGLIN 203

1897 qtpflmdrcmp-----mgistglidsqikase-----flgyweprlarlinggsyn 1946

204 AWT-----AAENDRWPMIOINLORKMRVGVITOGAKRIGSEYIKSVKIYVNSDGTWAM 259

1947 awsvklaaeafaskpwydmqkevilgicqgakhylksctetelyvasnqimwql 2006

260 YKVGKTEEMVFRGNVDNNTPYANSEFPPIKAQYVRLYPOICRRHCTLRMELLCESLGC 319

2007 fkgstrnmvmyfngnsdastikengfdppivarylrirsptraynrplrllelqgcevgnc 2066

320 SEPLGKSGHIDYOYTASSVFRILNMDFTWEPKARLDKOGVNAWTSCHNDQSQWLQ 379

2067 stpqlmengkienqkltaskfwgdy--wepfrlragqavnaagkannkqwl 2124

380 VDLVPTKVTGIIITOGADDFGVDPVSGYKLAYSNDGEHMAVHODEKORDKVFQGNFPDN 439

2125 fdllkfkktaltitlgckslseemyvkelytlhyseqvewkpyrlksmwdkifegntnt 2184

440 -DTHRKAVIDPPIYARFIRLPLMSWYGRITLASELLGC 476

Db 2185 kgvknfnpipisrfirvlpktnqstlrllelqgc 2221

RESULT 14

ID Y49564 standard; Protein: 2224 AA.

Y49564;

13-JAN-2000 (first entry)

Human lipoprotein lipase protein sequence.

Human: coding sequence polymorphism; vascular pathology gene; polymorphic site; phenotype correlation; forensic; paternity testing; medicine; genetic analysis; vascular disease.

Homo sapiens.

MO9950454-A2.

07-OCT-1999.

26-MAR-1999: 99WO-US06473.

01-APR-1998: 98US-0054272.

(WHEE) WHITEHEAD INST BIOMEDICAL RES.

Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

WPI: 1999-620066/53.

N-PSDB: 232183.

Determination of polymorphisms in genes, especially those identifying predisposition to vascular disease

Disclosure: Fig 27; 134pp; English.

232159 to 232194 represent reference alleles for specifically claimed nucleic acid sequences from the present invention which comprise polymorphic sites as given in a table in the specification, selected from 92 single nucleotide polymorphisms in which the nucleotide at the polymorphic site is different from a nucleotide at the same site in a reference allele. The nucleic acids, and primers and probes, are used to identify polymorphisms, which may predispose an individual to disease, especially a vascular disease. They can also be used in phenotype correlations, forensics, paternity testing, medicine or genetic analysis. Y49550 to Y49573 represent the proteins which correspond to some of the reference alleles.

Sequence 2224 AA;

Query Match 25.6%; Score 684; DB 20; Length 2224;

Best Local Similarity 42.4%; Pred. No. 4,3e-32; Matches 143; Conservative 55; Mismatches 123; Indels 16; Gaps 4;

144 ECPGEFMRNCOYKCSGHLGIEGIIISNOQTASSNHRALFGLQKWPYARLNKKGLIN 203

1897 qtpflmdrcmp-----mgistglidsqikase-----flgyweprlarlinggsyn 1946

204 AWT-----AAENDRWPMIOINLORKMRVGVITOGAKRIGSEYIKSVKIYVNSDGTWAM 259

1947 awsvklaaeafaskpwydmqkevilgicqgakhylksctetelyvasnqimwql 2006

260 YKVGKTEEMVFRGNVDNNTPYANSEFPPIKAQYVRLYPOICRRHCTLRMELLCESLGC 319

2007 fkgstrnmvmyfngnsdastikengfdppivarylrirsptraynrplrllelqgcevgnc 2066

320 SEPLGKSGHIDYOYTASSVFRILNMDFTWEPKARLDKOGVNAWTSCHNDQSQWLQ 379

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:11:26 ; Search time 179.67 Seconds
(without alignments)
429.733 Million cell updates/sec

Title: US-09-237-981-10
Perfect score: 2669
Sequence: 1 MKHLVAMLLVGLSLGPGF.....MSWGRITLNSLLGCAEE 480

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1009251 seqs, 160854530 residues
Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	2669	100.0	480	US-08-480-2298-10	Sequence 10, Appl
2	2669	100.0	480	US-08-659-2356-10	Sequence 10, Appl
3	2669	100.0	480	US-09-237-981-10	Sequence 10, Appl
4	2519.5	94.4	513	US-08-480-2298-14	Sequence 14, Appl
5	2519.5	94.4	513	US-08-659-2356-14	Sequence 14, Appl
6	2519.5	94.4	513	US-09-237-981-14	Sequence 14, Appl
7	2495.5	93.5	481	US-08-659-2358-29	Sequence 29, Appl
8	1645.5	61.7	321	US-08-480-2298-21	Sequence 21, Appl
9	1645.5	61.7	321	US-08-659-2358-21	Sequence 21, Appl
10	1645.5	61.7	321	US-09-237-981-21	Sequence 21, Appl

11	1603.5	60.1	449	US-60-207-315-404	Sequence 404, App
12	1401	52.5	362	US-60-207-315-523	Sequence 523, App
13	1400	52.5	311	US-60-230-435-1519	Sequence 1519, App
14	1242	46.5	221	US-09-237-981-29	Sequence 29, Appl
15	1168.5	43.8	463	US-09-582-340-4	Sequence 4, Appl
16	1023	38.3	387	US-09-582-340-2	Sequence 2, Appl
17	1004	37.6	203	US-60-233-644-77	Sequence 77, Appl
18	1001	37.5	185	US-60-230-435-1806	Sequence 1806, App
19	933.5	35.0	320	US-08-480-2298-20	Sequence 20, Appl
20	933.5	35.0	320	US-08-659-2358-20	Sequence 20, Appl
21	933.5	35.0	320	US-09-237-981-20	Sequence 20, Appl
22	869	25.8	2224	PCT-US96-03881A-14	Sequence 14, Appl
23	689	25.8	2224	US-08-410-488-14	Sequence 14, Appl
24	684	25.6	2224	US-09-054-272-38	Sequence 38, Appl
25	684	25.6	2228	US-60-208-020-1170	Sequence 1170, App
26	684	25.6	2228	US-60-209-043-221	Sequence 221, App
27	684	25.6	2228	US-60-213-172-167	Sequence 167, App
28	684	25.6	2228	US-60-213-173-167	Sequence 167, App
29	675	25.3	2319	PCT-US00-13541-6	Sequence 6, Appl
30	675	25.3	2319	PCT-US01-05076A-28	Sequence 28, Appl
31	675	25.3	2319	PCT-US97-11155-6	Sequence 6, Appl
32	675	25.3	2319	PCT-US99-05193-6	Sequence 6, Appl
33	675	25.3	2319	US-08-475-201-6	Sequence 6, Appl
34	675	25.3	2319	US-09-315-179-6	Sequence 6, Appl
35	675	25.3	2319	US-09-523-656-28	Sequence 28, Appl
36	660.5	24.7	2304	US-09-035-141-4	Sequence 4, Appl
37	660.5	24.7	2304	US-09-324-867-4	Sequence 4, Appl
38	647.5	24.3	1438	PCT-US99-29169-1	Sequence 1, Appl
39	647.5	24.3	1438	US-09-209-916-1	Sequence 1, Appl
40	647.5	24.3	1457	US-08-869-309-47	Sequence 47, Appl
41	647.5	24.3	1457	US-09-001-039A-47	Sequence 47, Appl
42	647.5	24.3	1457	US-09-001-039A-47	Sequence 47, Appl
43	647.5	24.3	1457	US-09-190-941-47	Sequence 47, Appl
44	647.5	24.3	1471	US-08-683-839A-3	Sequence 3, Appl
45	647.5	24.3	1661	US-09-243-532-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-480-2298-10
Sequence 10, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480, 2298
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28, 462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
 TELE: 66141 Pennie
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 480 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-229B-10

Query Match 100.0%; Score 2669; DB 8; Length 480;
 Best Local Similarity 100.0%; Pred. No. 7e-227;
 Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAAAMLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGFAGPNC 60
 DB 1 MKHLVAAAMLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGFAGPNC 60
 QY 61 SVEVADDEKPTISAGPCIPNCHNGTCEISEAYRGDTFYGVCPCRGFNGIHCOHNI 120
 DB 61 SVEVADDEKPTISAGPCIPNCHNGTCEISEAYRGDTFYGVCPCRGFNGIHCOHNI 120
 QY 121 NECEAPCRNGICITDLVANYSCCEGPEFGRNCOYKCSGHLIGIGIISNOQITASSNH 180
 DB 121 NECEAPCRNGICITDLVANYSCCEGPEFGRNCOYKCSGHLIGIGIISNOQITASSNH 180
 QY 181 RALFGLQKWPYARLKKGLINAMTAENDRPWQINLQKRMVTVITOGAKRIGSP 240
 DB 181 RALFGLQKWPYARLKKGLINAMTAENDRPWQINLQKRMVTVITOGAKRIGSP 240
 QY 241 EYIKSYKIAYVNDGKTAMTKVKTNEEMVFRGNVDNNTPVANSFTPIKAQYALYPOI 300
 DB 241 EYIKSYKIAYVNDGKTAMTKVKTNEEMVFRGNVDNNTPVANSFTPIKAQYALYPOI 300
 QY 301 CRRHCTLMELLGCELSGCEPLGMSGHIODYQITASSVFRTLMMDFTWEPKARLKD 360
 DB 301 CRRHCTLMELLGCELSGCEPLGMSGHIODYQITASSVFRTLMMDFTWEPKARLKD 360
 QY 361 QGKVNAMTSGHNDOSQWLQVDLVPYTKVGIITOGAKDFGHVQVGSYKLAAYSNDGEHM 420
 DB 361 QGKVNAMTSGHNDOSQWLQVDLVPYTKVGIITOGAKDFGHVQVGSYKLAAYSNDGEHM 420
 QY 421 VHODEKORKKVFQGNFNDTHRKNIIDPPIYARFRIILPMSWYGRITLRSLLGCAEE 480
 DB 421 VHODEKORKKVFQGNFNDTHRKNIIDPPIYARFRIILPMSWYGRITLRSLLGCAEE 480

RESULT 2
 US-08-659-235B-10
 Sequence 10: Application US/08659235B
 GENERAL INFORMATION:
 APPLICANT: Quertermous, Thomas
 APPLICANT: Hogan, Brigid
 APPLICANT: Snodgrass, H. Ralph
 APPLICANT: Zupancic, Thomas J.
 TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/659,235B

FILING DATE: 05-JUN-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Poissant, Brian M.
 REGISTRATION NUMBER: 28,462
 REFERENCE/DOCKET NUMBER: 8907-034
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELE: 66141 Pennie
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 480 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-659-235B-10

Query Match 100.0%; Score 2669; DB 10; Length 480;
 Best Local Similarity 100.0%; Pred. No. 7e-227;
 Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAAAMLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGFAGPNC 60
 DB 1 MKHLVAAAMLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGFAGPNC 60
 QY 61 SVEVADDEKPTISAGPCIPNCHNGTCEISEAYRGDTFYGVCPCRGFNGIHCOHNI 120
 DB 61 SVEVADDEKPTISAGPCIPNCHNGTCEISEAYRGDTFYGVCPCRGFNGIHCOHNI 120
 QY 121 NECEAPCRNGICITDLVANYSCCEGPEFGRNCOYKCSGHLIGIGIISNOQITASSNH 180
 DB 121 NECEAPCRNGICITDLVANYSCCEGPEFGRNCOYKCSGHLIGIGIISNOQITASSNH 180
 QY 181 RALFGLQKWPYARLKKGLINAMTAENDRPWQINLQKRMVTVITOGAKRIGSP 240
 DB 181 RALFGLQKWPYARLKKGLINAMTAENDRPWQINLQKRMVTVITOGAKRIGSP 240
 QY 241 EYIKSYKIAYVNDGKTAMTKVKTNEEMVFRGNVDNNTPVANSFTPIKAQYALYPOI 300
 DB 241 EYIKSYKIAYVNDGKTAMTKVKTNEEMVFRGNVDNNTPVANSFTPIKAQYALYPOI 300
 QY 301 CRRHCTLMELLGCELSGCEPLGMSGHIODYQITASSVFRTLMMDFTWEPKARLKD 360
 DB 301 CRRHCTLMELLGCELSGCEPLGMSGHIODYQITASSVFRTLMMDFTWEPKARLKD 360
 QY 361 QGKVNAMTSGHNDOSQWLQVDLVPYTKVGIITOGAKDFGHVQVGSYKLAAYSNDGEHM 420
 DB 361 QGKVNAMTSGHNDOSQWLQVDLVPYTKVGIITOGAKDFGHVQVGSYKLAAYSNDGEHM 420
 QY 421 VHODEKORKKVFQGNFNDTHRKNIIDPPIYARFRIILPMSWYGRITLRSLLGCAEE 480
 DB 421 VHODEKORKKVFQGNFNDTHRKNIIDPPIYARFRIILPMSWYGRITLRSLLGCAEE 480

RESULT 3
 US-09-237-981-10
 Sequence 10: Application US/09237981
 GENERAL INFORMATION:
 APPLICANT: Quertermous, Thomas
 APPLICANT: Hogan, Brigid
 APPLICANT: Snodgrass, H. Ralph
 APPLICANT: Zupancic, Thomas J.
 TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States

ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 03-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-10

Query Match 100.0%; Score 2669; DB 16; Length 480;
Best Local Similarity 100.0%; Pred. No. 7e-227;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 MHLVAMLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
1 MHLVAMLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
DB 61 SVEVASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFTIGYVCKPRFGNHCQHN 120
61 SVEVASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFTIGYVCKPRFGNHCQHN 120
DB 121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
DB 121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
DB 181 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 240
181 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 240
DB 181 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 240
181 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 240
DB 241 EYISYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 300
241 EYISYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 300
DB 241 EYISYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 300
241 EYISYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 300
DB 301 CRRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 360
301 CRRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 360
DB 301 CRRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 360
301 CRRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 360
DB 361 OGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 420
361 OGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 420
DB 361 OGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 420
361 OGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 420
DB 421 VHODEKORKDKVFOGNFNDNTHRKVNIDPPIYARFIRILPMSWYGRITLRSELLGCAGEE 480
421 VHODEKORKDKVFOGNFNDNTHRKVNIDPPIYARFIRILPMSWYGRITLRSELLGCAGEE 480
DB 421 VHODEKORKDKVFOGNFNDNTHRKVNIDPPIYARFIRILPMSWYGRITLRSELLGCAGEE 480
421 VHODEKORKDKVFOGNFNDNTHRKVNIDPPIYARFIRILPMSWYGRITLRSELLGCAGEE 480

RESULT 4
US-08-480-229B-14
Sequence 14, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid

APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229B-14

Query Match 94.4%; Score 2519.5; DB 8; Length 513;
Best Local Similarity 94.2%; Pred. No. 1.2e-213;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;
DB 1 MHLVAMLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
1 MHLVAMLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
DB 33 MRSVAVMLLVGLSLGVPOFGKGDICDPNCPENGICLPLGLAVGSEFSCCEPGEFAGPNC 92
33 MRSVAVMLLVGLSLGVPOFGKGDICDPNCPENGICLPLGLAVGSEFSCCEPGEFAGPNC 92
DB 61 SVEVASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFTIGYVCKPRFGNHCQHN 120
61 SVEVASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFTIGYVCKPRFGNHCQHN 120
DB 93 SVEVASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFTIGYVCKPRFGNHCQHN 152
93 SVEVASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFTIGYVCKPRFGNHCQHN 152
DB 121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
DB 121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
121 NECAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
DB 153 NECEVPECKNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 212
153 NECEVPECKNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 212
DB 181 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 239
181 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 239
DB 213 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 272
213 RALFGLQKWYPYARLNKKGLINAMTAENDRWPWIOINLQKRWVTGVITOGAKRIGSP 272
DB 240 PEYIKYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 299
240 PEYIKYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 299
DB 273 PEYIKYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 332
273 PEYIKYKIAVSNDRKTAMVYKGTNEEMVFRGVNDNNTPVANSFTPIKAQYVRLYPO 332
DB 300 CRRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 359
300 CRRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 359
DB 333 VCRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 392
333 VCRHCTLMELLGCELSGCESEPLGKMSGHIQDYQITASSVFTLNMDFWEPKARLDK 392
DB 360 KOGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 419
360 KOGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 419
DB 393 KOGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 452
393 KOGKVNAMTSGHNDOSQWLOVDLLVPTKYGTITOGAKDFGHVQVGSYKLAIVSNDGEHM 452
DB 420 MVHODEKORKDKVFOGNFNDNTHRKVNIDPPIYARFIRILPMSWYGRITLRSELLGCAGEE 479
420 MVHODEKORKDKVFOGNFNDNTHRKVNIDPPIYARFIRILPMSWYGRITLRSELLGCAGEE 479

Db 453 TVYDEKORDKVFGNFDNDTHRKNVIDPPIYARHRIILPMSWYGRITLASELCTEE 512
QY 480 E 480
Db 513 E 513

RESULT 5
US-08-659-235B-14
; Sequence 14, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quatermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-659-235B-14

Query Match 94.4%; Score 2519.5; DB 10; Length 513;
Best Local Similarity 94.2%; Pred. No. 1.2e-213;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 MKHLVAMLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPFGFAPKCS 60
Db 33 MKRSVAWMLLVGLSLGVPOFGKGDICDPNCEGICLPLGLAVGSFSCCEPFGFAPKCS 92
QY 61 SVEVVASDEKPTISAGFCIFNPGCHNGTCEISEAYRGDTFFIGYVCKCPRGFGNGICQHNH 120
Db 93 SVEVVASDEEPTISAGFCIFNPGCHNGTCEISEAYRGDTFFIGYVCKCPRGFGNGICQHNH 152
QY 121 NECEAEPCRNAGICTDLVANYSCCEGCEFGMRKCYKCGSHGIGTBSGIIISNOQITASSNH 180
Db 153 NECEVEPCRNAGICTDLVANYSCCEGCEFGMRKCYKCGSHGIGTBSGIIISNOQITASSNH 212
QY 181 RALEGLQKWPYARLKKGLINAWTAENDRWP-WIQINLQKMRVGVITOGAKRIGS 239
Db 213 RALEGLQKWPYARLKKGLINAWTAENDRKMWIOINLQKMRVGVITOGAKRIGS 272
QY 240 PEYIKSKIYISNDGKTWMTYKKGTEINEMVFRGNVDNNTPYANSFTTPPIKAQYRLPYQ 299

Db 273 PEYIKSKIYISNDGKTWMTYKKGTEINEMVFRGNVDNNTPYANSFTTPPIKAQYRLPYQ 332
QY 300 ICRRHCTLRKELLGCELSGCSERPLGKSGHIDYQITASSVPRILNMFMEPRKALD 359
Db 333 VCRHCTLRKELLGCELSGCSERPLGKSGHIDYQITASSVPRILNMFMEPRKALD 392
QY 360 KQKRVNMTSGHNDQSOVLQVDLVPYTKVTGIIITOGAKDFGHVQVGSYKLAYSNDGEHW 419
Db 393 KQKRVNMTSGHNDQSOVLQVDLVPYTKVTGIIITOGAKDFGHVQVGSYKLAYSNDGEHW 452
QY 420 MVYDEKORDKVFGNFDNDTHRKNVIDPPIYARHRIILPMSWYGRITLASELCTEE 479
Db 453 TVYDEKORDKVFGNFDNDTHRKNVIDPPIYARHRIILPMSWYGRITLASELCTEE 512
QY 480 E 480
Db 513 E 513

RESULT 6
US-09-237-981-14
; Sequence 14, Application US/09237981
; GENERAL INFORMATION:
; APPLICANT: Quatermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/237,981
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235
; FILING DATE: 05-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-237-981-14

Query Match 94.4%; Score 2519.5; DB 16; Length 513;
Best Local Similarity 94.2%; Pred. No. 1.2e-213;
Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 MKHLVAMLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPFGFAPKCS 60


```

Db 33 MKRSVAVMLLVGLSGVPGFGKDIDCPNCPENGICLPLGLAVSGFCECPDFTDPNCS 92
QY 61 SYVEVASDEBKPTSAQPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNT 120
Db 93 SYVEVASDEEPTSAQPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNT 152
QY 121 NECEAPBCRNNGICTDLVANYSCCEPGFMGRNCQYKCSGHLIEGGIISNOQTASSNH 180
Db 153 NECEVEPCRNNGICTDLVANYSCCEPGFMGRNCQYKCSGHLIEGGIISNOQTASSNH 212
QY 181 RALFGLQKMPYARLNKKGLINAMTAENDRMP-WIOINLORKMRYTGYITOGAKRIGS 239
Db 213 RALFGLQKMPYARLNKKGLINAMTAENDRMP-WIOINLORKMRYTGYITOGAKRIGS 272
QY 240 PEYIKFYKIAVSDGKTAMMYKVTNEEMVRGVDNNTFYANSFPPPIKAQIVRLYPQ 299
Db 273 PEYIKFYKIAVSDGKTAMMYKVTNEEMVRGVDNNTFYANSFPPPIKAQIVRLYPQ 332
QY 300 ICRHCTLRMELLGCELSGCSSEPLGKSGHIQDYQITASSVFRILNMDMFWBRKARLD 359
Db 333 VCRHCTLRMELLGCELSGCSSEPLGKSGHIQDYQITASSVFRILNMDMFWBRKARLD 392
QY 360 KQKRVNAMTSGHNDOSOMLOVDLVPKVTGIIITOGADPGHVQFVSGYKLAISNDGEHW 419
Db 393 KQKRVNAMTSGHNDOSOMLOVDLVPKVTGIIITOGADPGHVQFVSGYKLAISNDGEHW 452
QY 420 MVHODEKORKDKVQGNFNDTHRKNVIDPPIYARFIRILPMSWYGRITLSELGCAEE 479
Db 453 TVXODEKORKDKVQGNFNDTHRKNVIDPPIYARFIRILPMSWYGRITLSELGCTEE 512
QY 480 E 480
Db 513 E 513

```

RESULT 7
US-08-659-235B-29
Sequence 29, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 481 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235B-29

Query Match 93.5%; Score 2495.5; DR 10; Length 481;
Best Local Similarity 93.6%; Pred. No. 1,46-211;
Matches 450; Conservative 8; Mismatches 22; Indels 1; Gaps 1;

1 MKHVAAMLLVGLSGVPGFGKDIDCPNCPENGICLPLGLAVSGFCECPDFTDPNCS 60
1 MKRSVAVMLLVGLSGVPGFGKDIDCPNCPENGICLPLGLAVSGFCECPDFTDPNCS 60
Db 1 MKRSVAVMLLVGLSGVPGFGKDIDCPNCPENGICLPLGLAVSGFCECPDFTDPNCS 60
QY 61 SYVEVASDEBKPTSAQPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNT 120
Db 61 SYVEVASDEBKPTSAQPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNT 120
QY 61 SYVEVASDEBKPTSAQPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNT 120
Db 61 SYVEVASDEBKPTSAQPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNT 120
QY 121 NECEAPBCRNNGICTDLVANYSCCEPGFMGRNCQYKCSGHLIEGGIISNOQTASSNH 180
Db 121 NECEVEPCRNNGICTDLVANYSCCEPGFMGRNCQYKCSGHLIEGGIISNOQTASSNH 180
QY 181 RALFGLQKMPYARLNKKGLINAMTAENDRMP-WIOINLORKMRYTGYITOGAKRIGS 239
Db 181 RALFGLQKMPYARLNKKGLINAMTAENDRMP-WIOINLORKMRYTGYITOGAKRIGS 240
QY 240 PEYIKFYKIAVSDGKTAMMYKVTNEEMVRGVDNNTFYANSFPPPIKAQIVRLYPQ 299
Db 241 PEYIKFYKIAVSDGKTAMMYKVTNEEMVRGVDNNTFYANSFPPPIKAQIVRLYPQ 300
QY 300 ICRHCTLRMELLGCELSGCSSEPLGKSGHIQDYQITASSVFRILNMDMFWBRKARLD 359
Db 301 VCRHCTLRMELLGCELSGCSSEPLGKSGHIQDYQITASSVFRILNMDMFWBRKARLD 360
QY 360 KQKRVNAMTSGHNDOSOMLOVDLVPKVTGIIITOGADPGHVQFVSGYKLAISNDGEHW 419
Db 361 KQKRVNAMTSGHNDOSOMLOVDLVPKVTGIIITOGADPGHVQFVSGYKLAISNDGEHW 420
QY 420 MVHODEKORKDKVQGNFNDTHRKNVIDPPIYARFIRILPMSWYGRITLSELGCAEE 479
Db 421 TVXODEKORKDKVQGNFNDTHRKNVIDPPIYARFIRILPMSWYGRITLSELGCTEE 480
QY 480 E 480
Db 481 E 481

```

RESULT 8
US-08-480-229B-21
Sequence 21, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B

FILED DATE: 07-JUN-1995
CLASSIFICATION: 336
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-229B-21

Query Match 61.7%; Score 1645.5; DB 8; Length 321;
Best Local Similarity 95.0%; Pred. No. 8.6e-137;
Matches 304; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 158 CSGLGIEGIIISNOOITASSNHRALFGLQKWPYARLNKGLINMTAENDRW-PW1 216
DB 1 CSGLGIEGIIISNOOITASSNHRALFGLQKWPYARLNKGLINMTAENDRW1 60
QY 217 QINLQRMRYGVITOGAKRIGSPKIKSYKIAVNDGKTWAMKVKGTNEWVRGNVD 276
DB 61 QINLQRMRYGVITOGAKRIGSPKIKSYKIAVNDGKTWAMKVKGTNEWVRGNVD 120
QY 277 NNPYANSFTPIPKAOYVRLYPOICRHCTLMELLGELSGCSEPLGMSGHIQDYQT 336
DB 121 NNPYANSFTPIPKAOYVRLYPOICRHCTLMELLGELSGCSEPLGMSGHIQDYQT 180
QY 337 ASSVFRLLNMDFTWEPKRAKLDQKGVAMTSGHNDOSQMLQYDLVPTKVTGIIITOGA 396
DB 181 ASSVFRLLNMDFTWEPKRAKLDQKGVAMTSGHNDOSQMLQYDLVPTKVTGIIITOGA 240
QY 397 KDFGHVQFVGSYKLAISNDGEHMVHODEKQKRDVFOGNEFNDTHRKNVDPPIYARFI 456
DB 241 KDXGHVQFVGSYKLAISNDGEHMVHODEKQKRDVFOGNEFNDTHRKNVDPPIYARFI 300
QY 457 RILPMSWYGRITLSELGIC 476
DB 301 RILPMSWYGRITLSELGIC 320

RESULT 9

US-08-659-235B-21
Sequence 21, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quetermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/659,235B
FILED DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-659-235B-21

Query Match 61.7%; Score 1645.5; DB 10; Length 321;
Best Local Similarity 95.0%; Pred. No. 8.6e-137;
Matches 304; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 158 CSGLGIEGIIISNOOITASSNHRALFGLQKWPYARLNKGLINMTAENDRW-PW1 216
DB 1 CSGLGIEGIIISNOOITASSNHRALFGLQKWPYARLNKGLINMTAENDRW1 60
QY 217 QINLQRMRYGVITOGAKRIGSPKIKSYKIAVNDGKTWAMKVKGTNEWVRGNVD 276
DB 61 QINLQRMRYGVITOGAKRIGSPKIKSYKIAVNDGKTWAMKVKGTNEWVRGNVD 120
QY 277 NNPYANSFTPIPKAOYVRLYPOICRHCTLMELLGELSGCSEPLGMSGHIQDYQT 336
DB 121 NNPYANSFTPIPKAOYVRLYPOICRHCTLMELLGELSGCSEPLGMSGHIQDYQT 180
QY 337 ASSVFRLLNMDFTWEPKRAKLDQKGVAMTSGHNDOSQMLQYDLVPTKVTGIIITOGA 396
DB 181 ASSVFRLLNMDFTWEPKRAKLDQKGVAMTSGHNDOSQMLQYDLVPTKVTGIIITOGA 240
QY 397 KDFGHVQFVGSYKLAISNDGEHMVHODEKQKRDVFOGNEFNDTHRKNVDPPIYARFI 456
DB 241 KDXGHVQFVGSYKLAISNDGEHMVHODEKQKRDVFOGNEFNDTHRKNVDPPIYARFI 300
QY 457 RILPMSWYGRITLSELGIC 476
DB 301 RILPMSWYGRITLSELGIC 320

RESULT 10

US-09-237-981-21
Sequence 21, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quetermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/237, 981
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/659, 235
; FILING DATE: 05-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28, 462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 penile
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-237-981-21

```

Query Match 61.7%; Score 1645.5; DB 16; Length 321;

Best Local Similarity 93.0%; Pred. No. 8.6e-137;

Matches 304; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

```

QY 158 CSGLIEGIIISNOQITASSNHRALFGLQKWPYARLNKGLINAMTAENDRW-PWI 216
DB 1 CSPLGIEGIIISNOQITASSNHRALFGLQKWPYARLNKGLINAMTAENDRWNRWI 60
QY 217 QINLQKRMRYGTGTGAKRISPEYIKSYKIAVSNQKTMAMKVGTEEMVFRNDV 276
DB 61 QINLQKRMRYGTGTGAKRISPEYIKSYKIAVSNQKTMAMKVGTEEMVFRNDV 120
QY 277 NNTPYANSFTPIKAQYVRLYPQICRHCOTLRLMELLCGELSGCSEPLGMSGH 336
DB 121 NNTPYANSFTPIKAQYVRLYPQICRHCOTLRLMELLCGELSGCSEPLGMSGH 180
QY 337 ASSVFTLNMDMTPEPRKARLDKQKVNAMTSGHNDOSQMLQVLDLVPKVGITQGA 396
DB 181 ASSVFTLNMDMTPEPRKARLDKQKVNAMTSGHNDOSQMLQVLDLVPKVGITQGA 240
QY 397 KDFGHQVFGSYKLAAYSDNGEHMNVHODEKQKDKVFGNFDNDTTHRKNVTDPEIYARFI 456
DB 241 KDFGHQVFGSYKLAAYSDNGEHMNVHODEKQKDKVFGNFDNDTTHRKNVTDPEIYARFI 300
QY 457 RILPMSWYGRITLRSELLGC 476
DB 301 RILPMSWYGRITLRSELLGC 320

```

RESULT 11

US-60-207-315-404

; Sequence 404, Application US/60207315

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; FILE REFERENCE: CLO00601

; CURRENT APPLICATION NUMBER: US/60/207, 315

; CURRENT FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 528

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 404

; LENGTH: 449

; TYPE: PRT

; ORGANISM: HUMAN

; FEATURE:

; NAME/KEY: VARIANT

```

; LOCATION: (1)...(449)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-60-207-315-404

```

Query Match

Best Local Similarity 60.1%; Score 1603.5; DB 23; Length 449;

Matches 299; Conservative 7; Mismatches 7; Indels 119; Gaps 3;

```

QY 77 PCINPCHNGCTGEISEARGDPTGYCKGPRGNHCHQ----- 118
DB 2 PCINPCHNGCTGEISEARGDPTGYCKGPRGNHCHQ----- 118
QY 119 -----NINEGAEPKRNIGICTDLVANYSCGCPGEMGRNCOYKCSG 161
DB 122 RGLTGLQLLRPECGAPNINECEVPCNKGICTDLVANYSCGCPGEMGRNCOYKCSG 161
QY 162 LGIEGIIISNOQITASSNHRALFGLQKWPYARLNKGLINAMTAENDRWPMI----- 216
DB 182 LGIEGIIISNOQITASSNHRALFGLQKWPYARLNKGLINAMTAENDRWPMI----- 216
QY 217 -----QINLQKRMRYGTGTGAKRISPEYIKSYKIAVSNQKTMAMKVGTEEM 269
DB 242 CHLPQWQINLQKRMRYGTGTGAKRISPEYIKSYKIAVSNQKTMAMKVGTEEM 301
QY 270 VFRGNVNDNTPYANSFTPIKAQYVRLYPQICRHCOTLRLMELLCGELSGCSEPLGMSGH 329
DB 302 VFRGNVNDNTPYANSFTPIKAQYVRLYPQICRHCOTLRLMELLCGELSGCSEPLGMSGH 361
QY 330 IODVQITASSVFTLNMDMTPEPRKARLDKQKVNAMTSGHNDOSQMLQ----- 379
DB 362 IODVQITASSVFTLNMDMTPEPRKARLDKQKVNAMTSGHNDOSQMLQ----- 379
QY 380 --VDLVPKVT 389
DB 422 TALAMVPTAAT 433

```

RESULT 12

US-60-207-315-523

; Sequence 523, Application US/60207315

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; FILE REFERENCE: CLO00601

; CURRENT APPLICATION NUMBER: US/60/207, 315

; CURRENT FILING DATE: 2000-05-30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 523

; LENGTH: 362

; TYPE: PRT

; ORGANISM: HUMAN

; US-60-207-315-523

Query Match 52.5%; Score 1401; DB 23; Length 362;

Best Local Similarity 87.1%; Pred. No. 3.8e-115;

Matches 257; Conservative 8; Mismatches 6; Indels 24; Gaps 2;

```

QY 119 NINEGAEPKRNIGICTDLVANYSCGCPGEMGRNCOYKCSGHLIEGIIISNOQITASS 178
DB 52 DINECEVPCNKGICTDLVANYSCGCPGEMGRNCOYKCSGHLIEGIIISNOQITASS 111
QY 179 NHRALFGLQKWPYARLNKGLINAMTAENDRWPMI-----QINLQKRMRY 226
DB 112 THRALFGLQKWPYARLNKGLINAMTAENDRWPMI-----QINLQKRMRY 171

```

QY 227 TGVITOGAKRIGSPETIKSYKIAVSDGKTWAMYKVGTEMEVFRGVNDNTPYANSFT 286
 DB 172 TGVITOGAKRIGSPETIKSYKIAVSDGKTWAMYKVGTEMEVFRGVNDNTPYANSFT 231
 QY 287 PPKAQQYRLPYQCRHCHTLRMELLGCELSGCEPLGKSGHIDYOITASSVPTLNM 346
 DB 232 PPKAQQYRLPYQCRHCHTLRMELLGCELSGCEPLGKSGHIDYOITASSVPTLNM 291
 QY 347 DMFTWEPKARLDKQGVNMTSGHNDQSQWLQ-----VDLIVPTKVT 389
 DB 292 DMFTWEPKARLDKQGVNMTSGHNDQSQWLQPTLISVEVSGTALALMVPAT 346

RESULT 13
 US-60-230-435-1519
 ; Sequence 1519, Application US/60230435
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: CL000768
 ; CURRENT APPLICATION NUMBER: US/60/230,435
 ; CURRENT FILING DATE: 2000-09-06
 ; NUMBER OF SEQ ID NOS: 2991
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1519
 ; LENGTH: 311
 ; TYPE: PRT
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(311)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-60-230-435-1519

Query Match 52.5%; Score 1400; DB 23; Length 311;
 Best Local Similarity 87.4%; Pred. No. 3,8e-115;
 Matches 257; Conservative 7; Mismatches 6; Indels 24; Gaps 2;
 QY 120 INECVPCRNKGICTDLVANYSCCEPGEPMGRNCOYKSGHIGIEGIIISNOQITASSN 179
 DB 2 INECVPCRNKGICTDLVANYSCCEPGEPMGRNCOYKSGHIGIEGIIISNOQITASSN 61
 QY 180 HRALEGLQKWPYARLNKKGLINAMTAENDRMPWT-----QINLQKRMKVT 227
 DB 62 HRALEGLQKWPYARLNKKGLINAMTAENDRMPWTQELRGCHLGPMDQINLQKRMKVT 121
 QY 228 GVTIOGAKRIGSPETIKSYKIAVSDGKTWAMYKVGTEMEVFRGVNDNTPYANSFT 287
 DB 122 GVTIOGAKRIGSPETIKSYKIAVSDGKTWAMYKVGTEMEVFRGVNDNTPYANSFT 181
 QY 288 PPKAQQYRLPYQCRHCHTLRMELLGCELSGCEPLGKSGHIDYOITASSVPTLNM 347
 DB 182 PPKAQQYRLPYQCRHCHTLRMELLGCELSGCEPLGKSGHIDYOITASSVPTLNM 241
 QY 348 DMFTWEPKARLDKQGVNMTSGHNDQSQWLQ-----VDLIVPTKVT 389
 DB 242 DMFTWEPKARLDKQGVNMTSGHNDQSQWLQPTLISVEVSGTALALMVPAT 295

RESULT 14
 US-09-237-981-29
 ; Sequence 29, Application US/09237981
 ; GENERAL INFORMATION:
 ; APPLICANT: Quentemus, Thomas
 ; APPLICANT: Hogan, Bridgid
 ; APPLICANT: Snodgrass, H. Ralph
 ; APPLICANT: Zupancic, Thomas J.
 ; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
 ; TITLE OF INVENTION: CELL LOCUS-1

NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/237,981
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/659,235
 ; FILING DATE: 05-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Poissant, Brian M.
 ; REGISTRATION NUMBER: 28,462
 ; REFERENCE/DOCKET NUMBER: 8907-0034-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 Pennie
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 221 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal
 US-09-237-981-29

Query Match 46.5%; Score 1242; DB 16; Length 221;
 Best Local Similarity 98.2%; Pred. No. 2,1e-101;
 Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKHLYAAMLVGLSLGVDFGKGDIDNPNCENGICISGLADDSFSCCEPGRGAPMCS 60
 DB 1 MKHLYAAMLVGLSLGVDFGKGDIDNPNCENGICISGLADDSFSCCEPGRGAPMCS 60
 QY 61 SVEVYASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFIFYVCKPGRFGNGIHCQHN 120
 DB 61 SVEVYASDEEKPISAGPCIPNCHNGTCEISEAYRGDTFIFYVCKPGRFGNGIHCQHN 120
 QY 121 NECEAPCRNKGICTDLVANYSCCEPGEPMGRNCOYKSGHIGIEGIIISNOQITASSN 180
 DB 121 NECEAPCRNKGICTDLVANYSCCEPGEPMGRNCOYKSGHIGIEGIIISNOQITASSN 180
 QY 181 RALFGLQKWPYARLNKKGLINAMTAENDRMPWTQINL 220
 DB 181 RALFGLQKWPYARLNKKGLINAMTAENDRMPWTQIVT 220

RESULT 15
 US-09-582-340-4
 ; Sequence 4, Application US/09582340
 ; GENERAL INFORMATION:
 ; APPLICANT: INSERM
 ; APPLICANT: INSTITUT CURIE
 ; APPLICANT: CNRS
 ; TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants
 ; TITLE OF INVENTION: thereof
 ; FILE REFERENCE: Lactadherin
 ; CURRENT APPLICATION NUMBER: US/09/582,340
 ; CURRENT FILING DATE: 2000-09-15
 ; NUMBER OF SEQ ID NOS: 6


```

OY      12  G5L5VPOFGKDJONPNPNCENGIGTCL5GLADDSF-----SCCEGEGFAGMPCSSV- EVASD 68
Db      133  G5FIREIKRGECSQNTTPPSVYKSPGPEPEKPSLECTIYVAPRMSIIIEFESE 192
OY      69  EKPRTSAPCIPNPNCHNGTCEISEANRGDF-----IGYV--KCP---RGNGIHCQ 117
Db      193  DLEDSNP-----GSMFCYDRLEIWEFDPDVPHIGRQOGQTPRIRSSGILSM 245
OY      118  HNINECEAEFCRNGICITDVLVANS-CEGGEFMRNCQYKSGHIGIEGIIISMOIT 177
Db      246  --VFYDSDAIARKEG-----FSANSVLQSSVE-----DFKMEALGEMSEIHSDIT 29
OY      176  ASSNHRALRLOLQWYRYVARLKKGLNAMPAAENDMWPQIOINLOKRRARYGVYTOGA- 234
Db      293  ASSQYST-----NMSAENSRLNYP--NGMTPEGDSYREMIQVLDGLRFTYAVGTQCAI 345
OY      235  -KRIGSEYIKSYKLAYSNDGQTKWAMYKVGCTNEEMVFRGNDNNTPYANSFTPIKAQY 29

```

Db 346 SKETRRKRYVYKIDYSSNGEDW--ITTEKGNKPYLEQGNTPIDVVAVVAPPLITRF 403
Qy 294 VRLYPOICRRHCTLMELLCESLG--CSEPLGKSGHIODYQIYASSVFETLMMDFTW 351
Db 404 VRIKPAWEMGISMRFEVYCKITDPCSGMLGWSGLISDQITSS-----NQGDRNW 457
Qy 352 EPRKARL--DKOGKVNAWT---SGHNDOSOMLOVDLLVPTKVTGTTTGADGPGHVPVG 406
Db 458 MPENIRLVTSRSG---WALPPAPHSYINEMLOIDGEEKYVIGITIDGKHREKVFMR 513
Qy 407 SYKLAYSNDGEHMMVHODEKORDKVFOGNFNDTHRRKNVIDPPIYAFRILP-WSWYG 465
Db 514 KFKIGYSNNGSDMKMIMDDSKRAKSFEGNNNDYDPELTRF-PALSTFRIRIYPERATHG 572
Qy 466 RITLRSELGCAEE 479
Db 573 GLGLRMELLCGEVE 586

RESULT 2
PCT-US00-35017A-1285
; Sequence 1285, Application PC/TUS0035017A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: PCT/US00/35017A
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US09/468,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1285
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-35017A-1285

Query Match 17.8%; Score 475; DB 1; Length 957;
Best Local Similarity 31.4%; Pred No. 3.1e-35;
Matches 155; Conservative 66; Mismatches 207; Indels 66; Gaps 21;

Qy 12 GLSLGVPOFGKIDICNPNCENGICLSGLADDSF--SCECPGAPGNCSSVY-EVASD 68
Db 167 GFSIRYEIIFKRGPCSQNYTPPGVIRKSPGFEKYPNSLECTYIVFAFKMSEIILDFESE 226
Qy 69 EEPKPSAGPCIPNCHNGGTCSEAYRGDTF-----IGVVC--KCP---RGFNGIHQ 117
Db 227 DLEFDSNP-----GGMFCRYDLEIMDFPDVGPVHGRYCGKRTGRIRSSGILSM 279
Qy 118 HNINCEAEPCRNCGICITDLYANYS--ECPGEFMRNCQYKCSGHLGIEGITSNOIT 175
Db 280 --VEYTSALAKEG-----FSANYSVLOSYSSE-----DFKCMALMESGEIHSOIT 326
Qy 176 ASSNHRALFGLQKWPYRATLANKGLINAMTAENDRMPIODINORCMRTAGVITOGA- 234
Db 327 ASSQYST-----NMSAERSLNPPE--NGMTGEGDSYRREMIOVDGLRFTYATGCAI 379
Qy 235 -KRIGSPETIKSYKIAVSDGKTWAMKYKGTNEMKVRGNVNDNTPYANSFTPIKAY 293
Db 380 SKETRRKRYVYKIDYSSNGEDW--ITTEKGNKPYLEQGNTPIDVVAVVAPPLITRF 437
Qy 294 VRLYPOICRRHCTLMELLCESLG--CSEPLGKSGHIODYQIYASSVFETLMMDFTW 351
Db 438 VRIKPAWEMGISMRFEVYCKITDPCSGMLGWSGLISDQITSS-----NQGDRNW 491
Qy 352 EPRKARL--DKOGKVNAWT---SGHNDOSOMLOVDLLVPTKVTGTTTGADGPGHVPVG 406
Db 407 MPENIRLVTSRSG---WALPPAPHSYINEMLOIDGEEKYVIGITIDGKHREKVFMR 513
Qy 466 RITLRSELGCAEE 479
Db 573 GLGLRMELLCGEVE 586

Db 548 KFKIGYSNNGSDMKMIMDDSKRAKSFEGNNNDYDPELTRF-PALSTFRIRIYPERATHG 606
Qy 466 RITLRSELGCAEE 479
Db 607 GLGLRMELLCGEVE 620

RESULT 3
US-09-583-638-4
; Sequence 4, Application US/09583638
; GENERAL INFORMATION:
; APPLICANT: KLAGSBRUN, MICHAEL
; APPLICANT: SOKER, SHAY
; APPLICANT: MIAO, HUA-OUAN
; APPLICANT: TAKASHIMA, SEIJI
; TITLE OF INVENTION: NEUROPTILINS AND USE THEREOF IN METHODS FOR
; FILE REFERENCE: 701039-48800
; CURRENT APPLICATION NUMBER: US/09/583,638
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: PCT/US98/26127
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 60/069,155
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 60/069,687
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-583-638-4

Query Match 16.9%; Score 450.5; DB 5; Length 931;
Best Local Similarity 34.9%; Pred. No. 4.9e-33;
Matches 118; Conservative 61; Mismatches 122; Indels 37; Gaps 14;

Qy 156 YKSGHGIGGGIISNOITASSNHRALFGLQKWPYRATLANKGLINAMTAENDMWP 215
Db 275 FQCVNPLGMSGRANQISAST-----YSDGRWTPQOSRLH--GDDNGWTPNDSKEY 328
Qy 216 IQINDQRKMYTGVITOGA--KRIGSPETIKSYKIAVSDGKTWAMKYKGTNEMVFRG 273
Db 329 LOYDLRFLMTALATOGAISRETONGYVYKYLEVSTNGEDMWVYR-HGKNRK-VFOA 386
Qy 274 NVDNNTPYANSFTPIKAYVRLYPOICRRHCTLMELLCESLG--CSEPLGKSGHIQ 331
Db 387 NNDATEVYVNLKLAFLITRFVRIKRTPTWHSIGALALELFCRCVYDAPCSNMLGSLIA 446
Qy 332 DYQITASSVFETLMMDFTWEPKARL--DKOG---KVNAMTSGHNDOSOMLOVDLLVPT 386
Db 447 DSOISAST-----OYLWSPSARLVSRSRGWFPRIPOAPG---EKLQVLDLCTPK 496
Qy 387 KVTGIIITOGAKDGHV-----OFVGSYKLAYSNDGEHMMVHODEKORDKVFOGNFND 440
Db 497 TVKGVITOGARGDSTIYAEARAFVAKFKYSYSLNGKMEYIODPRQGRPLFEGNMHYD 556
Qy 441 THRRKNVIDPPIYAFRILP--WSWYGRITLRSELG 476
Db 557 TPDIRRFD-PIDPOYVAVYPERWSPAG-IGMRLEVLGC 592

RESULT 4
US-09-828-366-21
; Sequence 21, Application US/09828366
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.

ORGANISM: mouse
US-09-467-997-1

Query Match 9.9%; Score 264.5; DB 5; Length 1964;
Best Local Similarity 35.2%; Pred. No. 7.8e-16;
Matches 44; Conservative 18; Mismatches 42; Indels 21; Gaps 3;

QY 31 CENGICLSGLADDSFSCPEGFAGPNCSSVVEVASDEKPTSGPCIPNCHNGTCE 90
DB 897 CONGGLCID--TGSSYRCRCPGFQGLCOD-----NVPCEPNCHHGSTCV 942
QY 91 ISEAYRBDITGYCKPRPFNGIHCOHINECEAPCRNGICTDLVANYSCPEPGE 150
DB 943 POPP-----GYVCCAPGEGONCSKYLDACQSPCHNHGTCISRGPGFCACPGFV 995
QY 151 GRNCQ 155
DB 996 GLRCE 1000

RESULT 8
PCT-US01-09226-65
Sequence 65, Application PC/TUS0109226
GENERAL INFORMATION:
APPLICANT: SMITHKLINE BEECHAM CORPORATION
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP50018
CURRENT APPLICATION NUMBER: PCT/US01/09226
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 60/192,158
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/192,668
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: 60/200,166
PRIOR FILING DATE: 2000-04-27
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 65
LENGTH: 734
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-09226-65

Query Match 9.3%; Score 249; DB 1; Length 734;
Best Local Similarity 34.9%; Pred. No. 5.5e-15;
Matches 59; Conservative 25; Mismatches 69; Indels 16; Gaps 4;

QY 317 SCGEPELGKMSGHIOYQITASSVFRLNMDTWPRKARLDKQKV-----NATWS 369
DB 113 TGC-PPUGLESLRVSDBRLASS-----SQSFLGPHKGRNLITQSGLDGDLYDQAWKA 165
QY 370 GHNDQSOMLQVDLVPYKVTGIIITGAKDFGHVQVGYKLAYSNGEHMWHODEKOR 429
DB 166 EEDQADAMPVODAGHPFRSGVITGGRNSYWRVDWTSYKVOFSNDSRMTWSRNHSSGM 225
QY 430 DKVQGFNDTTHKKNYIDPPIYARFIRILPMGMY--GRITLSELGCG 476
DB 226 DAVEPANSDEPVALNLLPEQVAFRTIRLLPOTWLOGGAPCLAEITLAC 274

RESULT 9
PCT-US01-11988-770
Sequence 770, Application PC/TUS0111988
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: PCT/US01/11988
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/229, 358

PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 770
LENGTH: 383
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-11988-770

Query Match 9.0%; Score 241; DB 1; Length 383;
Best Local Similarity 36.1%; Pred. No. 1.3e-14;
Matches 48; Conservative 18; Mismatches 47; Indels 20; Gaps 6;

QY 25 IC-NPNCENGICLSGLADDSFSCPEGFAGPNCSSVVEVASDEKPTSGPC--IPN 81
DB 94 ICTQSPQNGGQCMD--GGGEYHCYCLPFGHGRDC-----ERK--AGPEQNGS 140
QY 82 PCHNGTCEISEAYRGDTFIGYCKPRPFNGIHCOHINECEAPCRNGICTDLVANY 141
DB 141 PCRNAGGCGQDDGFA---LNFTRCLYGVGARCEVNVDDCLMRCAAGATCIGINRF 196
QY 142 SCECPGEFGMRC 154
DB 197 SCLCPGEFGMRC 209

RESULT 10
US-09-833-245-770
Sequence 770, Application US/09833245
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229, 358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256, 931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199, 384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 770
LENGTH: 383
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-245-770

Query Match 9.0%; Score 241; DB 5; Length 383;
Best Local Similarity 36.1%; Pred. No. 1.3e-14;
Matches 48; Conservative 18; Mismatches 47; Indels 20; Gaps 6;

QY 25 IC-NPNCENGICLSGLADDSFSCPEGFAGPNCSSVVEVASDEKPTSGPC--IPN 81
DB 94 ICTQSPQNGGQCMD--GGGEYHCYCLPFGHGRDC-----ERK--AGPEQNGS 140
QY 82 PCHNGTCEISEAYRGDTFIGYCKPRPFNGIHCOHINECEAPCRNGICTDLVANY 141
DB 141 PCRNAGGCGQDDGFA---LNFTRCLYGVGARCEVNVDDCLMRCAAGATCIGINRF 196
QY 142 SCECPGEFGMRC 154
DB 197 SCLCPGEFGMRC 209

RESULT 11

TYPE: PRI
ORGANISM: Homo sapiens
PCT-US01-10484-63

Search completed: May 23, 2001, 06:13:52
Job time: 146 sec

Query Match 7.3%; Score 195.5; DB 1; Length 3014;

Best Local Similarity 28.8%; Pred. No. 2.3e-09;

Matches 45; Conservative 16; Mismatches 52; Indels 43; Gaps 6;

QY 24 DIGNPNCENGICISGLADDS-----FSCCEPGEFAGPNCSSVY 63

DB 1305 NICLRPCCENYMKCVSLRFDSSAPLSSTVLFPRHPINGLRKRCPPGFTGDCETEI 1364

QY 64 EVASDEKPTISAGPCIPNCHNGTCETISEAYRGDTFIGVCKCPRGFNGIHCOHNI--N 121

DB 1365 DL-----CYSDFCGANGRCRSREG-----GYTCECFEDFTGEHCEVDARSG 1405

QY 122 ECEAEPCRNNGICTD-LVANYSCCE-PGEFMRNCQ 155

DB 1406 RCANGVCKNGTCVNLIGGFHCVCPPGEYERPYCE 1441

RESULT 15

PCT-US01-01302-90

Sequence 90, Application PC/TUS0101302

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc., et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PTZ07PCT

CURRENT APPLICATION NUMBER: PCT/US01/01302

CURRENT FILING DATE: 2001-01-17

NUMBER OF SEQ ID NOS: 192

SOFTWARE: Patent Ver. 2.0

SEQ ID NO 90

LENGTH: 213

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (1)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (17)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (179)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (195)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (205)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

PCT-US01-01302-90

Query Match 7.3%; Score 194; DB 1; Length 213;

Best Local Similarity 27.3%; Pred. No. 1e-10;

Matches 42; Conservative 19; Mismatches 69; Indels 24; Gaps 3;

QY 9 LVGSLGVPOFGKGDICN-----PNCENGICISGLADDSFSCCEPGEFAGPNCSS 61

DB 35 LFEYLSFSLIYIKNSFCNKNTKRLSNCQNNSTCKDFSKDNCSDPTANNLDKDCDN 94

QY 62 YVEVASDEKPTISAGPCIPNCHNGTCETISEAYRGDTFIGVCKCPRGFNGIHCOHNI 121

DB 95 MKD-----PCFSNFCQGSATCVNTPGER-----SFLCKCPPEYSGTIOETTIG 137

QY 122 ECEAEPCRNNGICTDLVANYSCCEPGEFMRNCQ 155

DB 138 SCGKNSCGHGICHODPIYPVCICPAGYAGRCE 171

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:11:26 ; Search time 38.64 Seconds
(without alignments)
238.644 Million cell updates/sec

Title: US-09-237-981-10
Perfect score: 2669
Sequence: 1 MKHLVAMLVGLSLGVPQGF.....MSWTGRTLLSELGCAEE 480

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2669	100.0	480	US-08-480-229C-10	Sequence 10, Appl
2	2669	100.0	480	US-08-659-235C-10	Sequence 10, Appl
3	2519.5	94.4	513	US-08-480-229C-14	Sequence 14, Appl
4	2519.5	94.4	513	US-08-659-235C-14	Sequence 14, Appl
5	1645.5	61.7	321	US-08-480-229C-21	Sequence 21, Appl
6	1645.5	61.7	321	US-08-659-235C-21	Sequence 21, Appl
7	1242	46.5	221	US-08-480-229C-29	Sequence 29, Appl
8	1242	46.5	221	US-08-659-235C-29	Sequence 29, Appl
9	1160.5	43.5	463	US-08-162-402B-9	Sequence 9, Appl
10	1136	38.3	387	US-08-162-402B-8	Sequence 8, Appl
11	1023	35.0	320	US-08-480-229C-20	Sequence 20, Appl
12	933.5	35.0	320	US-08-659-235C-20	Sequence 20, Appl
13	933.5	35.0	320	US-08-659-235C-20	Sequence 20, Appl
14	694	26.0	2183	US-08-746-111-5	Sequence 5, Appl
15	675	25.3	2319	US-08-212-133A-8	Sequence 8, Appl
16	675	25.3	2319	US-08-474-503-6	Sequence 6, Appl
17	675	25.3	2319	US-08-670-707A-6	Sequence 6, Appl
18	675	25.3	2319	US-09-037-601-6	Sequence 6, Appl
19	675	25.3	2319	PCT-US94-13200-6	Sequence 6, Appl
20	651.5	24.4	2332	US-08-276-594A-2	Sequence 2, Appl
21	648.5	24.3	2351	5422260-1	Sequence 2, Appl
22	647.5	24.3	1471	US-08-683-839B-3	Sequence 3, Appl
23	647.5	24.3	1661	US-08-882-083-2	Sequence 2, Appl
24	647.5	24.3	1661	US-08-558-107-2	Sequence 2, Appl
25	647.5	24.3	1661	US-09-243-539-2	Sequence 2, Appl
26	647.5	24.3	2332	US-07-864-004B-4	Sequence 4, Appl
27	647.5	24.3	2332	US-08-251-937A-4	Sequence 4, Appl

28	647.5	24.3	2332	1	US-08-212-133A-2	Sequence 2, Appl
29	647.5	24.3	2332	1	US-08-474-503-2	Sequence 2, Appl
30	647.5	24.3	2332	2	US-08-670-707A-2	Sequence 2, Appl
31	647.5	24.3	2332	4	US-09-037-601-2	Sequence 4, Appl
32	647.5	24.3	2332	5	PCT-US93-03275-4	Sequence 2, Appl
33	647.5	24.3	2332	5	PCT-US94-13200-2	Sequence 2, Appl
34	647.5	24.3	2351	1	US-08-121-202-2	Sequence 2, Appl
35	647.5	24.3	2351	1	US-08-366-851A-2	Sequence 2, Appl
36	641.5	24.0	2351	6	5171844-2	Sequence 3, Appl
37	639	23.9	217	1	US-07-607-538C-3	Sequence 3, Appl
38	639	23.9	217	2	US-08-162-402B-3	Sequence 2, Appl
39	639	23.9	218	1	US-07-607-538C-2	Sequence 2, Appl
40	639	23.9	218	2	US-08-162-402B-2	Sequence 2, Appl
41	636.5	23.8	1443	4	US-08-670-707A-39	Sequence 39, Appl
42	636.5	23.8	1443	4	US-09-037-601-39	Sequence 39, Appl
43	636.5	23.8	2133	2	US-08-670-707A-37	Sequence 37, Appl
44	636.5	23.8	2133	4	US-09-037-601-37	Sequence 4, Appl
45	502	18.8	218	1	US-07-607-538C-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-480-229C-10
Sequence 10, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quertermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY REGULATED ENDOTHELIAL
NUMBER OF INVENTION: CELL LOCUS-1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480, 229C
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Polsant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229C-10

Query Match 100.0%; Score 2669; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 7e-213;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 MKHLVAMLVGLSLGVPQGFICNPCENGIGLISGLADSFSCCECFAPNCS 60

```

Db 1 MKHLVAMMLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
QY 61 SVEVASDEEKPISAGCIPNCPCHNGTCISBAYRGDTFTIGYVCKCPRGFNGIHCOHNI 120
Db 61 SVEVASDEEKPISAGCIPNCPCHNGTCISBAYRGDTFTIGYVCKCPRGFNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180
Db 121 NECEAEPCRNNGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180
QY 181 RALFGLQKWTPYYARLNKGLINAMTAENDRPMWIOINLQKRMVYGVITOGAKRIGSP 240
Db 181 RALFGLQKWTPYYARLNKGLINAMTAENDRPMWIOINLQKRMVYGVITOGAKRIGSP 240
QY 241 EYIKSYKIAVSDNGKTAMKVKGTNEEMVFRGNVDNNTPYANSFTPPKAOYVRLYPOI 300
Db 241 EYIKSYKIAVSDNGKTAMKVKGTNEEMVFRGNVDNNTPYANSFTPPKAOYVRLYPOI 300
QY 301 CRRHCTLRMELLCGELSGCEPGLGMSGHIQDYQITASSVFRTLMDMTWEPKARLDK 360
Db 301 CRRHCTLRMELLCGELSGCEPGLGMSGHIQDYQITASSVFRTLMDMTWEPKARLDK 360
QY 361 QGKVNAMTSGHNDOSOMLYVDLVPTRKVTGIITOGAKDGHQVFGSYKLAISNGEHHM 420
Db 361 QGKVNAMTSGHNDOSOMLYVDLVPTRKVTGIITOGAKDGHQVFGSYKLAISNGEHHM 420
QY 421 VHODEKQRKDVQGNFNDNTHRKNVIDPPIYARFIRILPMSWYGRITLRSLLCAEE 480
Db 421 VHODEKQRKDVQGNFNDNTHRKNVIDPPIYARFIRILPMSWYGRITLRSLLCAEE 480

```

RESULT 2

US-08-659-235C-10
Sequence 10, Application US/08659235C
Patent No. 5877281

GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennle
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
Topology: linear

MOLECULE TYPE: protein
US-08-659-235C-10

Query Match 100.0%; Score 2669; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 7e-213;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MKHLVAMMLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
Db 1 MKHLVAMMLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCPCPGFAGPNC 60
QY 61 SVEVASDEEKPISAGCIPNCPCHNGTCISBAYRGDTFTIGYVCKCPRGFNGIHCOHNI 120
Db 61 SVEVASDEEKPISAGCIPNCPCHNGTCISBAYRGDTFTIGYVCKCPRGFNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180
Db 121 NECEAEPCRNNGICTDLVANYSCCEPGEFGRNCQYKCSGHLGIEGGIISNOQITASSNH 180
QY 181 RALFGLQKWTPYYARLNKGLINAMTAENDRPMWIOINLQKRMVYGVITOGAKRIGSP 240
Db 181 RALFGLQKWTPYYARLNKGLINAMTAENDRPMWIOINLQKRMVYGVITOGAKRIGSP 240
QY 241 EYIKSYKIAVSDNGKTAMKVKGTNEEMVFRGNVDNNTPYANSFTPPKAOYVRLYPOI 300
Db 241 EYIKSYKIAVSDNGKTAMKVKGTNEEMVFRGNVDNNTPYANSFTPPKAOYVRLYPOI 300
QY 301 CRRHCTLRMELLCGELSGCEPGLGMSGHIQDYQITASSVFRTLMDMTWEPKARLDK 360
Db 301 CRRHCTLRMELLCGELSGCEPGLGMSGHIQDYQITASSVFRTLMDMTWEPKARLDK 360
QY 361 QGKVNAMTSGHNDOSOMLYVDLVPTRKVTGIITOGAKDGHQVFGSYKLAISNGEHHM 420
Db 361 QGKVNAMTSGHNDOSOMLYVDLVPTRKVTGIITOGAKDGHQVFGSYKLAISNGEHHM 420
QY 421 VHODEKQRKDVQGNFNDNTHRKNVIDPPIYARFIRILPMSWYGRITLRSLLCAEE 480
Db 421 VHODEKQRKDVQGNFNDNTHRKNVIDPPIYARFIRILPMSWYGRITLRSLLCAEE 480

```

RESULT 3

US-08-480-229C-14
Sequence 14, Application US/08480229C
Patent No. 5874562

GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEFAX: 66141 Pennie
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 513 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-229C-14

Query Match 94.4%; Score 2519.5; DB 2; Length 513;
 Best Local Similarity 94.2%; Pred. No. 1.7e-200;
 Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 MKHLVAMLLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCCEPFGAGNCS 60
 DB 33 MKRSVAWMLVGLSLGVPQFGKGDICDPNCPENGICLPLGLAVGSFSCCPGFTDPNCS 92
 QY 61 SVEVASDEEKPVSAGPCIPNCPCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 120
 DB 93 SVEVASDEEKPVSAGPCIPNCPCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 152
 QY 121 NECEAEPCRNNGICTDLVANYSCCEPFGMNCQYKCSGHLGIEGIIISNOQITASSNH 180
 DB 153 NECEAEPCRNNGICTDLVANYSCCEPFGMNCQYKCSGHLGIEGIIISNOQITASSNH 212
 QY 181 RALEGLQKWYFYARLNKKGLINAMTAENDRMP-WIOINLQKKMYGVITQGAKRIGS 239
 DB 213 RALEGLQKWYFYARLNKKGLINAMTAENDRMP-WIOINLQKKMYGVITQGAKRIGS 272
 QY 240 PEYIKSYLAISNDGKTWAMYKVGTEEMVFRGVNDNTPYANSTFPPIKAQYVRLPYQ 299
 DB 273 PEYIKSYLAISNDGKTWAMYKVGTEEMVFRGVNDNTPYANSTFPPIKAQYVRLPYQ 332
 QY 300 ICRHCTLRMELLCGELSGCEPLGKMSGHIDYOITASSVFTLNMDFWEPKARLD 359
 DB 333 VCRHCTLRMELLCGELSGCEPLGKMSGHIDYOITASSVFTLNMDFWEPKARLD 392
 QY 360 KQGVNANWISGHNDOSQWLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGBHW 419
 DB 393 KQGVNANWISGHNDOSQWLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGBHW 452
 QY 420 MVHODEKORKDKVFOGNDNTHRKNIIDPPIYARIRILPMSWYGRITLRSSELLGCAEE 479
 DB 453 TYVODEKORKDKVFOGNDNTHRKNIIDPPIYARIRILPMSWYGRITLRSSELLGCTEE 512
 QY 480 E 480
 DB 513 E 513

RESULT 4
 US-08-659-235C-14
 ; Sequence 14, Application US/08659235C
 ; Patent No. 5877281

GENERAL INFORMATION:
 APPLICANT: Quartermours, Thomas
 APPLICANT: Hogan, Bridgid
 APPLICANT: Snodgrass, H. Ralph
 APPLICANT: Zupancic, Thomas J.
 TITLE OF INVENTION: DEVELOPMENTALLY REGULATED ENDOTHELIAL
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States

ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/659,235C
 FILING DATE: 05-JUN-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Poissant, Brian M.
 REGISTRATION NUMBER: 28,462
 REFERENCE/DOCKET NUMBER: 8907-0034-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEFAX: 66141 Pennie
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 513 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-659-235C-14

Query Match 94.4%; Score 2519.5; DB 2; Length 513;
 Best Local Similarity 94.2%; Pred. No. 1.7e-200;
 Matches 453; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 MKHLVAMLLVGLSLGVPQFGKGDICNPNCENGICLSGLADDSFSCCEPFGAGNCS 60
 DB 33 MKRSVAWMLVGLSLGVPQFGKGDICDPNCPENGICLPLGLAVGSFSCCPGFTDPNCS 92
 QY 61 SVEVASDEEKPVSAGPCIPNCPCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 120
 DB 93 SVEVASDEEKPVSAGPCIPNCPCHNGTCEISEAYRGDTFYGVCPCPGFNGIHCOHNI 152
 QY 121 NECEAEPCRNNGICTDLVANYSCCEPFGMNCQYKCSGHLGIEGIIISNOQITASSNH 180
 DB 153 NECEAEPCRNNGICTDLVANYSCCEPFGMNCQYKCSGHLGIEGIIISNOQITASSNH 212
 QY 181 RALEGLQKWYFYARLNKKGLINAMTAENDRMP-WIOINLQKKMYGVITQGAKRIGS 239
 DB 213 RALEGLQKWYFYARLNKKGLINAMTAENDRMP-WIOINLQKKMYGVITQGAKRIGS 272
 QY 240 PEYIKSYLAISNDGKTWAMYKVGTEEMVFRGVNDNTPYANSTFPPIKAQYVRLPYQ 299
 DB 273 PEYIKSYLAISNDGKTWAMYKVGTEEMVFRGVNDNTPYANSTFPPIKAQYVRLPYQ 332
 QY 300 ICRHCTLRMELLCGELSGCEPLGKMSGHIDYOITASSVFTLNMDFWEPKARLD 359
 DB 333 VCRHCTLRMELLCGELSGCEPLGKMSGHIDYOITASSVFTLNMDFWEPKARLD 392
 QY 360 KQGVNANWISGHNDOSQWLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGBHW 419
 DB 393 KQGVNANWISGHNDOSQWLOVDLVPYKVTGIIITOGAKDFGHVQFVGSYKLAISNDGBHW 452
 QY 420 MVHODEKORKDKVFOGNDNTHRKNIIDPPIYARIRILPMSWYGRITLRSSELLGCAEE 479
 DB 453 TYVODEKORKDKVFOGNDNTHRKNIIDPPIYARIRILPMSWYGRITLRSSELLGCTEE 512
 QY 480 E 480
 DB 513 E 513

RESULT 5
 US-08-480-229C-21
 ; Sequence 21, Application US/08480229C
 ; Patent No. 5674562


```

RESULT 7
US-08-480-229C-29
; Sequence 29, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quartermou, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Polissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-480-229C-29

Query Match 46.5%; Score 1242; DB 2; Length 221:
Best Local Similarity 98.2%; Pred. No. 2.3e-95;
Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKHLVAAMLLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGEFAGPNC 60
DB 1 MKHLVAAMLLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGEFAGPNC 60
QY 61 SVVEVASDEEKPPTSAGPCIPNCPCHNGGTCEISEAVRGDTFIGYVCKCPRGFNGIHCOHNI 120
DB 61 SVVEVASDEEKPPTSAGPCIPNCPCHNGGTCEISEAVRGDTFIGYVCKCPRGFNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNH 180
QY 181 RALFGLQKWYPYARLANKKGLINAMTAENDRMPWIOITV 220
DB 181 RALFGLQKWYPYARLANKKGLINAMTAENDRMPWIOITV 220

```

```

; APPLICANT: Quartermou, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Polissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-659-235C-29

Query Match 46.5%; Score 1242; DB 2; Length 221:
Best Local Similarity 98.2%; Pred. No. 2.3e-95;
Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKHLVAAMLLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGEFAGPNC 60
DB 1 MKHLVAAMLLVGLSLGVPOFGKGDICNPNCENGICISGLADDSFSCCEPGEFAGPNC 60
QY 61 SVVEVASDEEKPPTSAGPCIPNCPCHNGGTCEISEAVRGDTFIGYVCKCPRGFNGIHCOHNI 120
DB 61 SVVEVASDEEKPPTSAGPCIPNCPCHNGGTCEISEAVRGDTFIGYVCKCPRGFNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNH 180
QY 181 RALFGLQKWYPYARLANKKGLINAMTAENDRMPWIOITV 220
DB 181 RALFGLQKWYPYARLANKKGLINAMTAENDRMPWIOITV 220

RESULT 9
US-08-162-402B-9
; Sequence 9, Application US/08162402B
; Patent No. 5972337
; GENERAL INFORMATION:
; APPLICANT: CERIANI, ROBERTO L.
; APPLICANT: PETERSON, JERRY A.
; APPLICANT: LARROCCA, DAVID J.
; TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
; TITLE OF INVENTION: GLOBULE (HMBG) ANTIGEN, FRAGMENTS & FUSION PROTEIN

```

```

: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pretty, Schroeder & Poplawski
: STREET: 444 South Flower St., 19th Floor
: CITY: Los Angeles
: STATE: CA
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/162,402B
: FILING DATE: 03-DEC-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Amzel, Viviana
: REGISTRATION NUMBER: 30,930
: REFERENCE/DOCKET NUMBER: P66 38215
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 213-622-7700
: TELEFAX: 213-489-4210
: TELEX:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 463 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
: US-08-162-402B-9

Query Match      43.5%; Score 1160.5; DB 2; Length 463;
Best Local Similarity 48.5%; Pred. No. 3.2e-88;
Matches 222; Conservative 74; Mismatches 139; Indels 23; Gaps 8;

QY 23 GDI CNPNPCENGICLSGLADSFSCCEPCEGFAFGPCSSVVEASDEEPTNAGCIPMP 82
DB 25 GDCDSSICLNGSTCLTG-QDNNDIYCLCEGFTGLVCNE-----TERGPCSPMP 72
QY 83 CHNGTCEIS-EAYRGDTFIVGCKCPREFNGIHCOHININECEAEPCRNAGCICITLVANY 141
DB 73 CYNDACKLVLTIDRGDITFEYICCPVGYSGIHCETETNYNMD---GEYFTTAVPMT 129
QY 142 SCECPGEF--MGRNCOYKCSGHLGIEGIIISNOQITASSNHRALFGLQKMPYVARLNKK 199
DB 130 AVPTPAPPTDLSNNLNASRSTOLGMEGAIADSOISASYVMGFOLQWGPBLARLYRT 189
QY 200 GLINATAENBRWPWIOINLQKMKRVGTITOGAKRIGSPREYIKSYKAYSNDOCKTAM 259
DB 190 GIYANNAHSAINTSLPWIQVNLKMKVSGVMTOGASRAGRALYLTFTFYAAYSLDRKKEF 249
QY 260 YK-VKGTNEEMVFRGNVNDNTPYANSFTPIKAQVRLYPOICRHRCTLRMLLCECELSG 318
DB 250 IDDESGDKRE--FLGNLDNNSLKVNMENPTLEAEYIRLYPVSCHNGCITREFLLGCELHG 307
QY 319 CSEPLGAKSGHIODVOITASSVFRILNMDMTWEPKRAKLDKOGVNAWTSGHNDOSQL 378
DB 308 CLEPGLKKNNTIPDSQMSASSSYKTMNLRAPGMYPHLGLHLDNOGKI NAWTAQSNBAKBL 367
QY 379 QVDLAVPTVVTGIIIOGADFGHVOFVGSYKLAYSNDGEMHVMVHODEKORCKVQSGNFD 438
DB 368 QVDLGRQRYVTGIIIOGADFGHIOVESYKVAHSDDGVMQTVY--EEOGSSKVFQGMLD 425
QY 439 NDTHRKNVLDIPYARFIRILPMSWYGRITLRSLLGC 476
DB 426 NNSHKKNIFEPKPFMAKRVAVLPVSNHNRITLLELLGC 463
```

```

RESULT 10
US-08-162-402B-8
: Sequence 8, Application US/08162402B
: Patent No. 5972337
: GENERAL INFORMATION:
: APPLICANT: CERIANI, ROBERTO L.
: APPLICANT: PETERSON, JERRY A.
: APPLICANT: LAROCCA, DAVID J.
: TITLE OF INVENTION: 46 KALTON HUMAN MILK FAT
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pretty, Schroeder & Poplawski
: STREET: 444 South Flower St., 19th Floor
: CITY: Los Angeles
: STATE: CA
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/162,402B
: FILING DATE: 03-DEC-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Amzel, Viviana
: REGISTRATION NUMBER: 30,930
: REFERENCE/DOCKET NUMBER: P66 38215
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 213-622-7700
: TELEFAX: 213-489-4210
: TELEX:
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 465 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: peptide
: US-08-162-402B-8

Query Match      42.6%; Score 1136; DB 2; Length 465;
Best Local Similarity 47.8%; Pred. No. 3.4e-86;
Matches 228; Conservative 64; Mismatches 155; Indels 30; Gaps 8;

QY 4 LVAAMLVGLSVGPQFGKGDICNPNCENGICLSGLADSFSCCEPCEGFAFGPCSSVY 63
DB 15 LCASLLVAID-----DFCDSSICLNGSTCLTG-QDNNDIYCLCEGFTGLVCNE-- 62
QY 64 EVASDEEKPASGCPINPCNNGTGC-EISRAYRGDTFIVGCKCPREFNGIHCOHINE 122
DB 63 -----TERGICSNPNCHNGSLCEISOEYRGDVFPSYTCCLKAGVGNICEETETNY 113
QY 123 CEAPPCRNAGCICITLVANYSCECPGEF--MGRNCOYKCSGHLGIEGIIISNOQITASSNH 180
DB 114 YNLD---GEYFTTAVPNTAVPTPAPPTDLSNNLNASKCEPLCMENGNINANSIASSVA 170
QY 181 RALFGLQKWPYVARLKKGLINAWTAENDRWPIQINLQKMKRVGTITOGAKRIGSP 240
DB 171 VTFGLQHWVPELARLRKAGVNAWNPSSNDNDPWIOVNLRRMWYTGVTYQASRLASH 230
QY 241 EYISYKLAYSNCKTW-AMYKVKGTNEEMVFRGNVNDNTPYANSFTPIKAQVRLYLPQ 299
DB 231 EYLAERVAISLNGHEDFTIDVKKKHKEFV--GNMKNAVHVNLEFTEPVEAQVRLYLP 288
```

```

Y 197 AKKGLINMATAENDRMPIQINLQKKMRYGVITTOGAKRISPEYIKSYKATYNSDGT 256
Y 109 NAGAVNATPSSNDNPNPIQVNLRLRMVGTGVTOGASRLASHHEYLAFAKVAISLNGE 168
Y 257 W-AMKVKOTSEMEVRCGVNDNTPYANSEFPPIKAQVRYLPO-CRRHCTLRMELGCE 315
Db 169 FDFIHDVKKKHEEV--GWNKNNAVHVNLEFPVEAOYRLPISCHRACTRLFELGCE 226
Y 316 LSGCEPIELGMSGHIQDYQITASSYFERTLNDMFTWEPREKALDKQGVNATSGNDQS 375
Db 227 LMGCANPGLKKNNSIPDKQITASSSYKTFWGLFESMNSYALDKQGFNMNVAGSYGND 286
Y 376 QMLQYDLVLPYVYITITTOGAKDFGHVQFVGSYKATYNSDGEHMNVHODEKQKRVQ 435
Y 287 QMLQYDLGSKSEVTEGIIITOGARFNSVQFVASYKATYNSDANSMTETODPRGSSKIPG 346
Y 436 NFDNDTHRKNVDPPIYARFIRLPMWGRITLSELLGC 476
Db 347 NMDNHSKKNLEFETPLIARYLRLVAMHNRALRLELLGC 387

RESULT 12
US-08-480-229C-20
: Sequence 20, Application US/08480229C
: Patent No. 5874562
: GENERAL INFORMATION:
: APPLICANT: Quetternous, Thomas
: APPLICANT: Hogan, Bridgid
: APPLICANT: Snodgrass, H. Ralph
: APPLICANT: Zupancic, Thomas J.
: TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
: TITLE OF INVENTION: CELL LOCUS-1
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: United States
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/480,229C
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Poissant, Brian M.
: REGISTRATION NUMBER: 28,462
: REFERENCE/DOCKET NUMBER: 8907-0026-999
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: TEXES: 66141 Pennie
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 320 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-480-229C-20

Query Match 35.08; Score 933.5; DB 2; Length 320;
Best Local Similarity 54.18; Pred. No. 1,2e-69;
Matches 173; Conservative 53; Mismatches 93; Indels 1; Gaps 1;

```

```

Db 1 CSTQLMEGGAIALDSQISASVYMGFMGLQRGPELRLRYRTGIYVNAHNASVYDQSKRWI 60
Qy 217 QINLQRMKRVYVITQAGAKRIGSPYIKSYKLAISNDGKTWAMYKVGKTEEMVFRGNVD 276
Db 61 QVNLKRMKRVSVMTQASRAGRAEYLKTFKVAVSLDGXKRFEFIQDESGDKREFLGND 120
Qy 277 NMTFVANSFPPIKAQVRLYPCQCRHCTLRMLLGCGLSCGSEPLMGSKGHIDQYIT 336
Db 121 NNSLKVMNFPTLEAOYIRLYPVSCHRGCTLRPELLGCELCLEPLGLKNTTIPDSQMS 180
Qy 337 ASSVFTLNMDFTEPRKARLDKQKVNAMTSGHNDOSQMLQVDLVPYKVTGIITQGA 396
Db 181 ASSSYKTWNLRAFQWPHILGRDNGKINAMTAQNSAKEMQLQVDLGTQRYVGTIITQGA 240
Qy 397 KDFGHVQVGSYKLAISNDGEHMMVHODEKQKDKVFGNFDNDTHRKVNIDPIYARFT 456
Db 241 RDFGHIOVYESYKVAHSDGVQMTVYXXEQQSSKVFQGNLDNNSHKNIPEKPFMARYV 300
Qy 457 RILPMSWGRITLRLSELLGC 476
Db 301 RVLPSWMNRITLRLLELGC 320

```

RESULT 13
US-08-659-235C-20
Sequence 20, Application US/08659235C
Patent No. 5877281

```

GENERAL INFORMATION:
APPLICANT: Quatermous, Thomas
APPLICANT: Hogan, Bridg
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-659-235C-20

```

Query Match 35.0%; Score 933.5; DB 2; Length 320;
Best Local Similarity 54.1%; Pred. No. 1.2e-69;
Matches 133; Conservative 53; Mismatches 93; Indels 1; Gaps 1;

```

Qy 158 CSGLGIEGGIISNQITASSNHRALFGLQKWPYARLNKKGILNMTAEND-RMPWI 216
Db 1 CSTQLMEGGAIALDSQISASVYMGFMGLQRGPELRLRYRTGIYVNAHNASVYDQSKRWI 60
Qy 217 QINLQRMKRVYVITQAGAKRIGSPYIKSYKLAISNDGKTWAMYKVGKTEEMVFRGNVD 276
Db 61 QVNLKRMKRVSVMTQASRAGRAEYLKTFKVAVSLDGXKRFEFIQDESGDKREFLGND 120
Qy 277 NMTFVANSFPPIKAQVRLYPCQCRHCTLRMLLGCGLSCGSEPLMGSKGHIDQYIT 336
Db 121 NNSLKVMNFPTLEAOYIRLYPVSCHRGCTLRPELLGCELCLEPLGLKNTTIPDSQMS 180
Qy 337 ASSVFTLNMDFTEPRKARLDKQKVNAMTSGHNDOSQMLQVDLVPYKVTGIITQGA 396
Db 181 ASSSYKTWNLRAFQWPHILGRDNGKINAMTAQNSAKEMQLQVDLGTQRYVGTIITQGA 240
Qy 397 KDFGHVQVGSYKLAISNDGEHMMVHODEKQKDKVFGNFDNDTHRKVNIDPIYARFT 456
Db 241 RDFGHIOVYESYKVAHSDGVQMTVYXXEQQSSKVFQGNLDNNSHKNIPEKPFMARYV 300
Qy 457 RILPMSWGRITLRLSELLGC 476
Db 301 RVLPSWMNRITLRLLELGC 320

```

RESULT 14
US-08-746-111-5
Sequence 5, Application US/08746111
Patent No. 6066778

```

GENERAL INFORMATION:
APPLICANT: Ginsburg, David
APPLICANT: Cui, Jisong
TITLE OF INVENTION: Compositions and Methods for Screening
TITLE OF INVENTION: Compounds for Anticoagulant Activity
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESS: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,111
FILING DATE: 06-NOV-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: 0M-02536
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2183 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-746-111-5

```

Query Match 26.0%; Score 694; DB 3; Length 2183;
Best Local Similarity 43.8%; Pred. No. 8.6e-49;
Matches 144; Conservative 51; Mismatches 112; Indels 22; Gaps 5;

```

Db 1865 ECKMPGLSTGVISDSQIKASEY-----LTWEPRLARLNAGSYNAMIETALDPL 1918
QY 213 WPTIOLNORRMRYGTITOGAKRISPEYIKS-----YIAISNDCKTAMKVGCTNE 267
Db 1919 KPMQVMOKEVYVTTGTOTGAK-----HYLKSCTTEFEVVAASSDQTNQIIRGSGKS 1973
QY 268 EMFERGVNDNTPYANSFTPIKAQVRLYPOICRRHCTIRMLLCELSGSEPIGMS 327
Db 1974 VMYTGNSDSGTIKENLDPIYARIYRIHPTKSYNRPILRLDQCEVNGCSTPLGLE 2033
QY 328 GHIDYQITASSVFRILNMDFTWEPKARLDKQKVNANTSGHNDQSGMLQVLLVPTK 387
Db 2034 GRIDQKITASSFRKSWMGDY--WEPRLARLNAGRYNMAQAKANNKQWLQVLLKIKK 2091
QY 388 VTGIIITGADFGHVOFVGSYKLAYSNDGEHVMVHODEKQKDKVPOGNDNDTHRKNY 447
Db 2092 VTALVITOGCSLSEMYKSYSTIYSDQYAMKPYRKOKSMVKTREGNSTGKHKNEF 2151
QY 448 DPLIYAFIRILPMSWYGRITLRSLLGC 476
Db 2152 NPILISFRIRIPIKTMOSTALRLLEFGC 2180

```

RESULT 15

```

US-08-212-133A-8
; Sequence 8, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Iollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2319 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; AUTHORS: Elder, F.

```

```

; AUTHORS: Iakich, D.
; AUTHORS: Gitschier, J.
; TITLE: Sequence of the Murine Factor VIII cDNA.
; Patent No. 5663060
; JOURNAL: Genomics
; VOLUME: 16
; PAGES: 374-379
; DATE: 1993
; RELEVANT RESIDUES IN SEQ ID NO: 8: FROM 1 TO 2319
US-08-212-133A-8

```

Query Match 25.3%; Score 675; DB 1: Length 2319;
 Best Local Similarity 44.6%; Pred 3.5e-47;
 Matches 154; Conservative 42; Mismatches 119; Indels 30; Gaps 6;

```

QY 137 LVANYSCECPGEFGMGRNCQYKCSGHLGIEGIIISNOQITASSNHRALPGLQWPPYARL 196
Db 2002 LVYSKQCCIP-----LGMASGSIRDFQITASGSHY-----GQMAPNLARL 2040
QY 197 NKKGLINMTAENDRMPIQINLQKKRVYGVITOGAKRIGSPYIKSYIAISNDGKT 256
Db 2041 HYSGSINMASTKE--PFSWIKVDLLAPMIYHGIKTQGAROKRFSLSLYISQFTIMYSLDCK 2098
QY 257 WAMYRVKGINEEMVPRGVNDNTPYANSFTPIKAQVRLYPOICRRHCTIRMLLCELSG 316
Db 2099 WLSYQNGSTGILAMVFFGVNDSSGIKHNSFNPIIARIYRIHPTKSYNRPILRLDQCEV 2158
QY 317 SCGSEPLGKSGHIDYQITASSVFRILNMDFTWEPKARLDKQKVNANTSGHNDQ 375
Db 2159 NSCSIPLGWESKVIDTQITASSYF---TNMFATWSPSOARLHLQGRTNAMRPQVNDPK 2214
QY 376 OMLQVLDLVPTKVTGIIITOGAKDFGHVOFVGSYKLAYSNDGEHVMVHODEKQKDKV 435
Db 2215 OMLQVLDLQKTKMKTGIIITOGAKSLFTSMFVKEFLISSQDGHWT--OILYNGKVVQ 2272
QY 436 NFDNDTHRKNYIDPIYAFIRILPMSWYGRITLRSLLGCAGEE 480
Db 2273 NQDSSTPMMNSLDPPLRLRYLRIHPQIWEHQIALRLLEITGGEAQ 2317

```

Search completed: May 23, 2001, 06:13:18
 Job time: 112 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:12:49 ; Search time 62.53 seconds
(without alignments)
242.888 Million cell updates/sec

Title: US-09-237-981-29
Perfect score: 1262

Sequence: 1 MKHLVAAMLVLGLSLGVPOF.....INAWTAENDRPWPIQVTVG 221

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 67:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	399.5	31.7	409	2 T11743	PP47 protein - pig
2	389	30.8	463	1 A36479	milk fat globule m
3	385	30.5	427	2 S74211	PAS-6/7 protein pr
4	373.5	29.6	427	2 JC4915	ags protein precu
5	363.5	28.8	401	2 S65138	glycoprotein anti
6	341	27.0	2352	2 T30201	Notch homology prot
7	324.5	25.7	2524	2 A35844	Notch protein - Af
8	321	25.4	2437	2 S42612	transmembrane prot
9	318.5	25.2	1064	2 A40136	fibropellin Ia-s
10	314.5	24.9	728	2 I50719	C-Delta-1 - chicke
11	313.5	24.8	570	2 A48836	fibropellin C prec
12	312.5	24.8	2531	2 S18188	notch protein homo
13	312.5	24.8	2531	2 A46019	Notch-1 protein -
14	308.5	24.4	2703	1 A24420	notch protein - fr
15	303	24.0	1203	2 A46019	Notch B protein -
16	299.5	23.7	2531	2 T31070	notch homology - se
17	298	23.6	1408	2 S16148	gene serrate prote
18	297	23.5	2471	2 A49128	cell-fate determin
19	293.5	23.2	2555	2 A40043	notch protein homo
20	291.5	23.1	722	2 I48324	DELTA-like 1 - mou
21	290.5	23.0	2321	2 S78349	notch3 protein - h
22	289.5	22.9	861	2 A48825	Notch homology Motc
23	289.5	22.9	2318	2 S45306	slit-1 protein hom
24	287.5	22.8	1531	2 T42218	slit-1 protein hom
25	286	22.7	1220	2 A56136	Notch protein pre
26	280.5	22.2	293	2 B26637	neurogenic repetit
27	280.5	22.2	2139	2 A56172	crumbs protein - f
28	278	22.0	1025	2 T42626	secreted leucine-r
29	274.5	21.8	473	2 A56175	adhesive plaque pr

30	271.5	21.5	832	2 A31246	neurogenic protein
31	271.5	21.5	833	2 S19087	gene Delta protein
32	271.5	21.5	880	2 S00670	neurogenic repetit
33	271	21.5	1687	2 T30176	EGF repeat transme
34	270	21.4	1469	2 B36655	slit protein 2 pre
35	270	21.4	1480	2 A36665	slit protein 1 pre
36	268	21.2	308	3 JC7125	epidermal growth f
37	266.5	21.1	1523	2 T13953	MEGF5 protein - ra
38	264.5	21.0	1964	2 T09059	notch4 - mouse
39	256	20.3	601	2 T22025	hypothetical prote
40	250.5	19.8	383	2 S53716	delta-like homeoti
41	247	19.6	1429	2 S06434	homeotic protein 1
42	245.5	19.5	1372	2 T25933	hypothetical prote
43	243	19.3	502	2 T20130	hypothetical prote
44	243	19.3	838	2 T20125	hypothetical prote
45	242.5	19.2	385	2 S53718	homeotic protein d

ALIGNMENTS

RESULT 1

T11743

PP47 protein - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: T11743

R:Enslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toe Biol. Reprod. 58, 1057-1064, 1998

A:Title: Molecular cloning and characterization of P47, a novel boar sperm-associated

A:Reference number: Z17325; MUID:96206817

A:Accession: T11743

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-409 <ENS>

A:Cross-references: EMBL:Y11683; NID:q2652927; PIDN:CAN/2379.1; PID:q2652928

A:Experimental source: testis

C:Function:

A:Description: may be involved in membrane remodeling and/or function as a zona pellu

F:Superfamily: unassigned EGF-related proteins; EGF homology

F:6-40/Domain: EGF homology <EGF>

Query Match 31.7%; Score 399.5; DB 2; Length 409;

Best local similarity 39.6%; Pred. No. 2.1e-23;

Matches 80; Conservative 20; Mismatches 47; Indels 55; Gaps 5;

QY 23 GDICNPENPCNGICLSGLAD---SFSCCEPGEFAGPNCSSVEVASDEKPTSGPCI 79

DB 3 GPCDSSSLINGTCL-LDOPQKPFHCLCPGFTGLICNE-----TEKGFCF 49

QY 80 PNCCHNGTCF-ISEAYRGDFIGYCKPRGFNGIHCHONTNECEAEBCRNGIGCTDLV 138

DB 50 PNCCHNDAECVLDHARDVLFETICKCPHGYTGH----- 86

QY 139 ANYSCCEPGEFMRNCQYCSGHLGIEGIIISNOQTASSNHRALFGLOKMPYAAALNK 198

DB 87 -----CEICNAPVIGMETGAIADFOISASSMHLGFMGJORMAPELARLR 131

QY 199 KGLINAWTAENDRPWPIQTV 220

DB 132 AGIVNMTASNDYRNPWIVNL 153

RESULT 2

A36479

milk fat globule membrane protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A36479

R:Stubbins, J.D.; Lekutis, C.; Singer, K.L.; But, A.; Yuzuki, D.; Srinivasan, U.; Parry Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990

A:Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the


```

Best Local Similarity 42.1%; Pred. No. 2,46-18;
Matches 61; Conservative 17; Mismatches 45; Indels 22; Gaps 4;

QY      26  CNMPCENGIGTGLGLADSFSCCEPGFAGPCNCSVVEVASDEKPPISAGPCIPINPCHN 85
      |  ||:|||||  ||:  ||:|||||  ||  ||  ||:|||||  |  ||:|||||  |
Db      720  CATOPCQNGGCTSGI--NSYNACAPAKTYGVCE-----TEISPVPNPCEN 765

QY      86  GGTCEISATYKGDPIFGVCCPKRFGNIGHQHNINCEAEPCRNIGICTDPLVANYSEC 145
      ||:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      766  GATCPQESADY-----LAAYCQCPREGFRGPTCATDINECVNSPKNGGGCTNLVPTQCTC 820

QY      146  PGEEMGRMCQY---KCSGHLIEGG 167
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      821  SQGFTGKDCDPTDIDDCSSNPCLNG 845

RESULT  7
A35844
Xotch protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 13-Aug-1999
C:Accession: A35844
R:Coffman, C.; Harris, W.; Kintner, C.
Science 249, 1438-1441, 1990
A:Title: Xotch, the Xenopus homolog of Drosophila notch.
A:Reference number: A35844; MUID:90385285
A:Accession: A35844
A:status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-2524 <COF>

```

[illegible]

```

QY      84  HNGGT-----CEISEAYRGDFIFG---YVCKCPGRFN 112
      ||
Db      955  KNGANGCTDCVNSYTCQCQGFSGIHCESNTPROTBSSECFNGGTICIDGINTFTCCPPEFT 1014
      ||
QY      113  GTHCOHNIIECAEPORNGICTDLVANTSCBEPGFMRNCQ----- 155
      ||:||||:||||:||||:||||:||||:
Db      1015  GSYCOHDHIECDSPCLNLCQDSDSYGYKTCPOGYTGLNCONLVRWCDSPPCKNGKGC 1074
      ||
QY      156  -----YKC---SGHLGI 164
      ||:| ||:|
Db      1075  WQTNNFYRCECKSGWTV 1092

RESULT      8
SA2612
transmembrane protein precursor - zebra fish
C:Species: Brachydanio rerio (zebra fish)
  at 100% boot above 20.000

```

R; Bierkamp, C.; Campos-Ortega, J.A.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:19:00 ; Search time 34.32 seconds

(without alignments)
220.585 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 1262
Sequence: 1 MKHLYAMWLVLGSLGVPQF.....INAMTAENDRMPQVTVG 221

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SWISSPROT_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	399.5	31.7	409	1	MEGM_PIG
2	389	30.8	463	1	MEGM_MOUSE
3	382.5	30.3	427	1	MEGM_BOVIN
4	373.5	29.6	427	1	MEGM_RAT
5	324.5	25.7	2524	1	NORC_XENLA
6	321	25.4	2437	1	NORC_BRARE
7	318.5	25.2	1064	1	FBP1_STRPU
8	313.5	24.8	570	1	FBP1_STRPU
9	312.5	24.8	2531	1	NTC1_MOUSE
10	307.5	24.4	714	1	DL1_RAT
11	302.5	24.0	2703	1	NORC_DROME
12	298	23.6	1408	1	SERR_DROME
13	297	23.5	723	1	DL1_HUMAN
14	292.5	23.2	2444	1	DL1_MOUSE
15	291.5	23.1	722	1	DL1_MOUSE
16	289.5	22.9	2318	1	DL1_MOUSE
17	285.5	22.6	589	1	DL1_MOUSE
18	281.5	22.3	592	1	DL1_MOUSE
19	280.5	22.2	2139	1	CRB_DROME
20	278.5	22.1	618	1	DL1_HUMAN
21	275.5	21.8	387	1	MEGM_HUMAN
22	274.5	21.8	473	1	FP2_MYGA
23	271.5	21.4	833	1	DL_DROME
24	270	21.4	1480	1	SLIT_DROME
25	262.5	20.8	1964	1	NTC4_MOUSE
26	256	20.3	4590	1	FATH_HUMAN
27	250.5	19.8	383	1	DLK_HUMAN
28	247	19.6	1429	1	DLK_MOUSE
29	242.5	19.2	5147	1	FAT_DROME
30	235	18.6	1295	1	GLP1_CAEEL
31	225.5	17.9	3562	1	PGCV_CHICK
32	201				
33					

34	199.5	15.8	379	1	WIF1_MOUSE	Q9WU01 mus musculus
35	198	15.7	1257	1	PGCN_RAT	P55067 rattus norv
36	197	15.6	1268	1	PGCN_MOUSE	P55066 mus musculus
37	195	15.5	4543	1	LRL1_CHICK	P98157 gallus gall
38	190.5	15.1	379	1	WIF1_HUMAN	Q9Y5W5 homo sapien
39	187	14.8	2201	1	TEMA_HUMAN	P24821 homo sapien
40	185.5	14.7	515	1	APX1_CAEEL	P41990 caenorhabd
41	183	14.5	3358	1	PGCV_MOUSE	Q62059 mus musculus
42	181.5	14.4	3051	1	YNX3_CAEEL	P34576 caenorhabd
43	181	14.3	862	1	PGCV_MACNE	Q28858 macaca neme
44	181	14.3	3396	1	PGCV_HUMAN	P13611 homo sapien
45	178.5	14.1	816	1	NEL_CHICK	Q90827 gallus gall

ALIGNMENTS

RESULT	1	STANDARD	PRT	409 AA.
MEGM_PIG				
ID	P79385			
AC	01-NOV-1997 (Rel. 35, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	LACTADHERIN (MILK FAT GLOBULE-EGF FACTOR 8) (MEG-E8) (MEGM) (SPERM SURFACE PROTEIN SP47) (PP47).			
GN	MEG8.			
OS	Eus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCB1			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	Tissue-testis:			
RA	Ensembl M.A.:			
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.			
CC	- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-BINDING PROTEIN.			
CC	- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.			
CC	- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND SPERMATOZOAN.			
CC	- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.			
CC	- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	EMBL: Y11683; CAA72379.1; ..			
DR	HSSP: P00740; 11YA			
DR	InterPro: IPR000421; ..			
DR	InterPro: IPR000561; ..			
DR	Pfam: PF00008; EGF_2.			
DR	Pfam: PF00754; F5_F8_TypeC; 2.			
DR	PROSITE: PS00022; EGF_1; 2.			
DR	PROSITE: PS01186; EGF_2; 2.			
DR	PROSITE: PS01285; FA58C_1; 2.			
DR	PROSITE: PS01286; FA58C_2; 2.			
KW	Glycoprotein; Repeat; EGF-like domain.			
FT	DOMAIN 2 41			EGF-LIKE 1.
FT	DOMAIN 44 88			EGF-LIKE 2.
FT	DOMAIN 91 247			F5/8 TYPE C.
FT	DOMAIN 252 409			F5/8 TYPE C.
FT	SITE 67 69			CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID 6 17			BY SIMILARITY.
FT	DISULFID 11 29			BY SIMILARITY.
FT	DISULFID 31 40			BY SIMILARITY.
FT	DISULFID 91 247			BY SIMILARITY.
FT	DISULFID 234 238			BY SIMILARITY.

FT DISULFID 252 409 BY SIMILARITY.
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 409 AA: 45725 MW: B0C07AF80029927A CRC64;

Query Match 31.7%; Score 399.5; DB 1; Length 409;
 Best Local Similarity 39.6%; Pred. No. 4, 1e-26;
 Matches 80; Conservative 20; Mismatches 47; Indels 55; Gaps 5;

QY 23 GICNPENPCENGICLSGLADDSFSCCEPGEFAGPNCSSVVEVADSKPTSGAPCIP 79
 DB 3 GFCOSSSLCNGTCLDOPQKPFQICPEGFTGLICNE-----TEKGPCE 49
 QY 80 PNPCHNGTCE-ISEAYRGDFIVGCKCPRGFNGIHQHNINCEAPRCNGICIDLV 138
 DB 50 PNPCHNDACEYIDAHNRDVFTEYICCPHGYTGR----- 86
 QY 139 ANYSCCEPGEFMRNCOYKCSGHLGIEGGIISNOQTASSNHRALFGLOKWPYYAALNK 198
 DB 87 -----CEICNAPLGMETGALADPOISASSMHLGFMGLQRMAPELARLHR 131
 QY 199 KGLINAWTAENDRWPMIOQTV 220
 DB 132 AGIVNAWTAASNDKPMIOQVNL 153

RESULT 2
 MFGM_MOUSE STANDARD: PRT: 463 AA.
 ID MFGM_MOUSE STANDARD: PRT: 463 AA.
 AC P21956; P97800; (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (MFGM)
 DE (SPERM SURFACE PROTEIN SP47) (MP47).
 GN MFGE8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 23-35.
 RC TISSUE-Mammary gland;
 RX MEDLINE=91046008; PubMed=2122462;
 RA Stubbbs J.D., Iekutiis C., Singer K.L., Bul A., Yuzuki D.,
 RA Strinivasan U., Parry G.;
 RT "CDNA cloning of a mouse mammary epithelial cell surface protein
 RT to factor VIII-like sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
 RN [2]
 RP SEQUENCE OF 23-463 FROM N.A.
 RC TISSUE-Testis;
 RA Ensslin M.A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. ZONA PELLUCIDA-
 CC BINDING PROTEIN.
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELL SURFACES AND
 CC SPERMATOZOAN.
 CC -1- DEVELOPMENTAL STAGE: MRNA EXPRESSION IS DETECTABLE IN MAMMARY
 CC -1- TISSUE FROM NONPREGNANT ANIMALS & MAXIMAL IN THE LACTATING GLAND.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M38337; AAA39534.1; -
 DR EMBL: Y11684; CAA72380.1; -
 DR PIR: A36479; A36479.
 DR HSSP: P00740; 11XA.
 DR MGD: MGI:102768; Mfge8.
 DR InterPro: IPR000421; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001438; -
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00754; F5_F8_Type_C_2.
 DR PRINTS: PR00010; EGFBL00-
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 KW Signal; Glycoprotein; Repeat; EGF-like domain; Milk.
 FT SIGNAL 1 22
 FT CHAIN 1 22
 FT DOMAIN 23 463
 FT DOMAIN 24 61
 FT DOMAIN 64 108
 FT DOMAIN 148 303
 FT DOMAIN 308 463
 FT SITE 87 89
 FT SITE 28 39
 FT DISULFID 28 39
 FT DISULFID 33 49
 FT DISULFID 51 60
 FT DISULFID 68 79
 FT DISULFID 73 96
 FT DISULFID 98 107
 FT DISULFID 148 303
 FT DISULFID 290 294
 FT DISULFID 308 463
 FT CARBOHYD 61 61
 FT CARBOHYD 266 266
 FT CARBOHYD 316 316
 FT CARBOHYD 426 426
 FT CONFLICT 30 30
 FT CONFLICT 35 35
 FT CONFLICT 110 147
 FT CONFLICT 168 168
 FT CONFLICT 196 196
 FT CONFLICT 309 309
 FT CONFLICT 395 395
 SQ SEQUENCE 463 AA: 51465 MW: D78B6C6EBBA724D CRC64;
 G (IN REF. 2).
 Y -> S (IN REF. 2).
 H -> T (IN REF. 2).
 L -> S (IN REF. 2).
 E -> A (IN REF. 2).
 S -> F (IN REF. 2).
 N -> D (IN REF. 1; AA SEQUENCE).
 ETNYNLDGEVMEFTYAVPTAVPTAPPPDLSNMLASR ->

Query Match 30.8%; Score 389; DB 1; Length 463;
 Best Local Similarity 38.8%; Pred. No. 3, 4e-25;
 Matches 78; Conservative 30; Mismatches 75; Indels 18; Gaps 5;

QY 23 GICNPENPCENGICLSGLADDSFSCCEPGEFAGPNCSSVVEVADSKPTSGAPCIP 82
 DB 25 GFCOSSSLCNGTCLTGQDNIDYCLCPGEFTGLICNE-----TERGCSNP 72
 QY 83 GNGGTCES-DAYRGDFIVGCKCPRGFNGIHQHNINCEAPRCNGICIDLVAN 141
 DB 73 CYNDKACLVLTDFQGDVFTEYICCPHGYTGR-----GEYMFYAVPT 129
 QY 142 SCECPGEF-MGRNCOYKCSGHLGIEGGIISNOQTASSNHRALFGLOKWPYYAALNK 199
 DB 130 AVPTAPPPDLSNMLASRSTQGLMEGALIDSOISASVYVGMGMGLQRMGPELARL 189
 QY 200 GLINAWTAENDRWPMIOQTV 220
 DB 190 GIVNAWTAASNDKPMIOQVNL 210

RESULT 3
 MFGM_BOVIN STANDARD: PRT: 427 AA.
 ID MFGM_BOVIN STANDARD: PRT: 427 AA.
 AC Q95114; Q27959; P79344;

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8)
 DE (MGP57/53) (PAS-6/PAS-7 GLYCOPROTEIN) (MFGM) (SPERM SURFACE PROTEIN
 DE SP47) (BP47) (COMPONENTS 15/16).
 GN MFG-E8.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
 RC STRAIN-HOLSTEIN; TISSUE-Mammary gland;
 RA Hvarregard J., Andersen M.H., Berglund L., Rasmussen J.T.,
 RA Petersen T.E.;
 RT "Characterization of glycoprotein PAS-6/7 from membranes of bovine
 RT milk fat globules.";
 RL Eur. J. Biochem. 240:628-636(1996).
 RN [2]
 RP SEQUENCE OF 18-427 FROM N.A.
 RC TISSUE-Mammary gland;
 RX MEDLINE=96125736; PubMed=8541316;
 RX Aoki N., Kishi M., Taniguchi Y., Adachi T., Nakamura R.,
 RA Matsuda T.;
 RT "Molecular cloning of glycoprotein antigens MGP57/53 recognized by
 RT monoclonal antibodies raised against bovine milk fat globule
 RT membrane.";
 RL Biochim. Biophys. Acta 1245:385-391(1995).
 RN [3]
 RP SEQUENCE OF 19-427 FROM N.A.
 RC TISSUE-Testis;
 RA Esselin M.A.;
 RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RL [4]
 RP SEQUENCE OF 140-146; 174-187; 233-246 AND 422-427.
 RC TISSUE-Milk;
 RX MEDLINE=93250576; PubMed=8485470;
 RA Mather I.H., Banghart L.R., Lane W.S.;
 RT "The major fat-globule membrane proteins, bovine components 15/16 and
 RT guinea-pig GP 55, are homologous to MGF-E8, a murine glycoprotein
 RT containing epidermal growth factor-like and factor V/VIII-like
 RT sequences.";
 RL Biochem. Mol. Biol. Int. 29:545-554(1993).
 CC -1- FUNCTION: PROBABLY ASSOCIATES WITH PHOSPHOLIPIDS ON THE SURFACE OF
 CC MAMMARY EPITHELIAL CELLS AND MILK FAT GLOBULES. ZONA PELLUCIDA-
 CC BINDING PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM
 CC LACKS 53 AMINO ACIDS WITHIN THE P5/8 TYPE C 1 DOMAIN.
 CC -1- TISSUE SPECIFICITY: MILK AND SPERMATOZOAN.
 CC -1- PTM: THE 2 O-LINKED GLYCANS CONSIST OF GAL, GLCNAC AND FUC, WITH
 CC PROBABLY FUC AS REDUCING TERMINAL SUGAR.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 P5/8 TYPE C DOMAINS.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sb.ch/announce/
 CC or send an email to license@isb-sb.ch).
 CC
 CC EMBL: X01895; CAA62997.1; -
 CC EMBL: S80643; AAB35894.2; -
 CC EMBL: Y11719; CAA72406.1; -
 CC HSSP: P00740; IIXA.
 CC InterPro: IPR000421; -
 CC InterPro: IPR000561; -
 CC Pfam: PF00008; EGF; 2.

DR Pfam: PF00754; F5_F8_type_C; 2.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01285; FAS8C_1; 2.
 DR PROSITE: PS01286; FAS8C_2; 2.
 KW Signal: glycoprotein; Milk; Repeat; EGF-like domain;
 KW Alternative splicing.
 FT SIGNAL 1
 FT CHAIN 19 427 LACTADHERIN.
 FT DOMAIN 20 59 EGF-LIKE 1.
 FT DOMAIN 62 106 EGF-LIKE 2.
 FT DOMAIN 109 265 F5/8 TYPE C 1.
 FT DOMAIN 270 427 F5/8 TYPE C 2.
 FT SITE 85 87 CELL ATTACHMENT SITE (POTENTIAL).
 FT DISULFID 24 35 BY SIMILARITY.
 FT DISULFID 29 47 BY SIMILARITY.
 FT DISULFID 49 58 BY SIMILARITY.
 FT DISULFID 66 77 BY SIMILARITY.
 FT DISULFID 71 94 BY SIMILARITY.
 FT DISULFID 96 105 BY SIMILARITY.
 FT DISULFID 109 265 BY SIMILARITY.
 FT DISULFID 252 256 BY SIMILARITY.
 FT DISULFID 270 427 BY SIMILARITY.
 FT CARBOHYD 27 27 O-LINKED (FUC. . .) (IN PAS-6).
 FT CARBOHYD 34 34 O-LINKED (FUC. . .) (IN PAS-7).
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (HYBRID) (IN PAS-6
 FT CARBOHYD 59 59 AND PAS-7).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (HIGH MANNOSE) (IN
 FT CARBOHYD 227 227 PAS-6).
 FT VARSPLIT 169 221 MISSING (IN SHORT ISOFORM).
 FT CONFLICT 19 19 A -> F (IN REF. 1).
 FT CONFLICT 28 28 L -> Q (IN REF. 1).
 SQ SEQUENCE 427 AA; 47411 MW; 4CBBE3A1DCAEB24 CRC64;
 Query Match 30.3%; Score 382.5; DB 1; Length 427;
 Best Local Similarity 36.6%; Pred. NO. 1.le-24;
 Matches 74; Conservative 28; Mismatches 45; Indels 55; Gaps 5;
 QY 23 GDICNPNCNGICISGLD-----SFCSECPAGNCSSVVEVASDEKPTASGPCI 79
 DB 21 GDFCSSLCHHGICL--LNEDRRPPFCICPEBFTGLTNE-----TEHPCF 67
 QY 80 PNPCHNGTCEIS-EAYGDTFTIGYCKCPRGFNGICOHNINECAEPCRNIGICTDVL 138
 DB 68 PNPCHNDACQVTDSDSHRGDVFYQICCKPLGYVGHCEET----- 108
 QY 139 ANYSCCEGEMRNCQYKCGHGLIEGLISNOOITASSHRLAFQIKKYPYALNK 198
 DB 109 -----CTSPLOMOTGALADSOISASSMHLGFMGLORWAPELARLHQ 149
 QY 199 KGLINAMTAENDRMPWIOVTV 220
 DB 150 TGIYVAMTSGNYDKRPWIOVNL 171
 RESULT 4
 MFGM_RAT
 ID MFGM_RAT STANDARD; PRT; 427 AA.
 AC P70490.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LACTADHERIN PRECURSOR (MILK FAT GLOBULE-EGF FACTOR 8) (MFG-E8) (O-
 DE ACETYL GD3 GANGLIOSIDE SYNTHASE) (AGS) (MFGM).
 GN MFG-E8 OR AGS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;

```

RX MEDLINE=96374422; PubMed=8780713;
RA Ogura K., Nara K., Matanabe Y., Kohno K., Tai T., Sanai Y.;
RT "Cloning and expression of cDNA for O-acetyltransferase of GD3
RT ganglioside.";
RL Biochem. Biophys. Res. Commun. 225:932-938(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN PHOSPHOLIPID BINDING. SEEMS TO
CC PARTICIPATE IN THE O-ACETYLATION OF GD3 GANGLIOSIDE SIALIC ACID.
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: SPLEEN, LUNG, HEART, BRAIN AND MUSCLE.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D84068; BAA12210.1; -.
DR HSP: P00740; IYX.
DR InterPro: IPR000421; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR001438; -.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00754; F5_FR_type_C; 2.
DR PRINTS: PR00010; EGFBL00D.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FA58C_1; 2.
DR PROSITE: PS01286; FA58C_2; 2.
DR Signal: Glycoprotein; Repeat; EGF-like domain; Milk.
KW SIGNAL.
FT CHAIN 1 22 POTENTIAL.
FT DOMAIN 23 427 LACTADHERIN.
FT DOMAIN 64 108 EGF-Like 1.
FT DOMAIN 111 267 EGF-Like 2.
FT DOMAIN 272 427 F5/8 TYPE C 1.
FT DOMAIN 28 39 F5/8 TYPE C 2.
FT DISULFID 33 49 BY SIMILARITY.
FT DISULFID 51 60 BY SIMILARITY.
FT DISULFID 68 79 BY SIMILARITY.
FT DISULFID 73 96 BY SIMILARITY.
FT DISULFID 98 107 BY SIMILARITY.
FT DISULFID 111 267 BY SIMILARITY.
FT DISULFID 254 258 BY SIMILARITY.
FT DISULFID 272 427 BY SIMILARITY.
FT SITE 87 89 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQUENCE 427 AA; 47413 MW; EA8C8631F3EB6047 CRC64;

```

```

Query Match 29.6%; Score 373.5; DB 1; Length 427;
Best Local Similarity 38.1%; Pred. No. 6e-24;
Matches 75; Conservative 23; Mismatches 48; Indels 51; Gaps 5;

```

```

OY 23 GDICNPNCENGIGICSLGADDSFCECEPESGAPNCSSVVEASDEKPTSGACPINP 82
DB 25 GDFCSISLCLNGGTCIMG-QNDIYICLCPEGTGLVCNE-----TEKSPCSPPN 72
OY 83 CHNGSTCISE-AVRGDFIFGYCKPRGFGHICQININCEAEPGRNGICIDLVANY 141
DB 73 CFHDKACLVETDPTGDFETETICQCPVGSYSHCE----- 108
OY 142 SCECEGFEFMGRNCOYKSGHIGIEGGLISNQOITASSNHRALFGLOKWPYVALINKGL 201
DB 109 -----LG-----CSTKRLGLEGAIALDSQISASSVYMGFMGLQRMGPFLANLYRTGI 154
OY 202 INAWTAENDRKPWIOV 218

```

```

DB 155 VNAWTAASYDSKRWIOV 171
:|||||: | |||||
RESULT 5
NOTCH_XENLA STANDARD; PRT; 2524 AA.
ID NOTCH_XENLA
AC P21783;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG PRECURSOR (NOTCH PROTEIN).
GN NOTCH.
OS Xenopus laevis (African clawed frog).
OC Amphibia; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90385285; PubMed=2402639;
RA Colman C., Harris W., Kintner C.;
RT "Notch, the Xenopus homolog of Drosophila notch.";
RL Science 249:1438-1441(1990).
RN [2]
RP REVISIONS TO 1759-1782.
RA Kintner C.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M33874; AAB02039.1; -.
DR PIR: A35844; A35844.
DR HSP: P00740; IYX.
DR InterPro: IPR000152; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR000800; -.
DR InterPro: IPR001438; -.
DR InterPro: IPR001881; -.
DR InterPro: IPR002110; -.
DR Pfam: PF00008; EGF; 36.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00010; EGFBL00D.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 23.
DR PROSITE: PS00022; EGF_1; 34.
DR PROSITE: PS01186; EGF_2; 29.
DR PROSITE: PS01187; EGF_CA; 21.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2524 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG.
FT DOMAIN 20 1728 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1729 1750 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1751 2524 POTENTIAL.
FT DOMAIN 20 57 EGF-Like 1.
FT DOMAIN 58 99 EGF-Like 2.
FT DOMAIN 102 140 EGF-Like 3.
FT DOMAIN 141 177 EGF-Like 4.

```

```
FT DOMAIN 179 215 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 217 254 EGF-LIKE 6,
FT DOMAIN 256 292 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 294 332 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 334 370 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 371 409 EGF-LIKE 10,
FT DOMAIN 411 449 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 451 487 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 489 525 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 527 563 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 565 600 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 602 638 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 640 675 EGF-LIKE 17,
FT DOMAIN 677 713 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 715 750 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 752 788 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 790 826 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 828 866 EGF-LIKE 22,
FT DOMAIN 868 904 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 906 942 EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 944 980 EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 982 1018 EGF-LIKE 26,
FT DOMAIN 1020 1056 EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1058 1094 EGF-LIKE 28,
FT DOMAIN 1096 1142 EGF-LIKE 29,
FT DOMAIN 1144 1180 EGF-LIKE 30,
FT DOMAIN 1182 1218 EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1220 1264 EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1266 1304 EGF-LIKE 33,
FT DOMAIN 1306 1346 EGF-LIKE 34,
FT DOMAIN 1347 1383 EGF-LIKE 35,
FT DOMAIN 1386 1424 EGF-LIKE 36,
FT DOMAIN 1441 1478 EGF-LIKE 37,
FT REPEAT 1479 1520 LIN/NOTCH 1,
FT REPEAT 1521 1560 LIN/NOTCH 2,
FT REPEAT 1560 1600 LIN/NOTCH 3,
FT REPEAT 1600 1640 ANK 1,
FT REPEAT 1640 1680 ANK 2,
FT REPEAT 1680 1720 ANK 3,
FT REPEAT 1720 1760 ANK 4,
FT REPEAT 1760 1800 ANK 5,
FT REPEAT 1800 1840 ANK 6,
FT REPEAT 1840 1880 BY SIMILARITY,
FT REPEAT 1880 1920 BY SIMILARITY,
FT REPEAT 1920 1960 BY SIMILARITY,
FT REPEAT 1960 2000 BY SIMILARITY,
FT REPEAT 2000 2040 BY SIMILARITY,
FT REPEAT 2040 2080 BY SIMILARITY,
FT REPEAT 2080 2120 BY SIMILARITY,
FT REPEAT 2120 2160 BY SIMILARITY,
FT REPEAT 2160 2200 BY SIMILARITY,
FT REPEAT 2200 2240 BY SIMILARITY,
FT REPEAT 2240 2280 BY SIMILARITY,
FT REPEAT 2280 2320 BY SIMILARITY,
FT REPEAT 2320 2360 BY SIMILARITY,
FT REPEAT 2360 2400 BY SIMILARITY,
FT REPEAT 2400 2440 BY SIMILARITY,
FT REPEAT 2440 2480 BY SIMILARITY,
FT REPEAT 2480 2520 BY SIMILARITY,
FT REPEAT 2520 2560 BY SIMILARITY,
FT REPEAT 2560 2600 BY SIMILARITY,
FT REPEAT 2600 2640 BY SIMILARITY,
FT REPEAT 2640 2680 BY SIMILARITY,
FT REPEAT 2680 2720 BY SIMILARITY,
FT REPEAT 2720 2760 BY SIMILARITY,
FT REPEAT 2760 2800 BY SIMILARITY,
FT REPEAT 2800 2840 BY SIMILARITY,
FT REPEAT 2840 2880 BY SIMILARITY,
FT REPEAT 2880 2920 BY SIMILARITY,
FT REPEAT 2920 2960 BY SIMILARITY,
FT REPEAT 2960 3000 BY SIMILARITY,
FT REPEAT 3000 3040 BY SIMILARITY,
FT REPEAT 3040 3080 BY SIMILARITY,
FT REPEAT 3080 3120 BY SIMILARITY,
FT REPEAT 3120 3160 BY SIMILARITY,
FT REPEAT 3160 3200 BY SIMILARITY,
FT REPEAT 3200 3240 BY SIMILARITY,
FT REPEAT 3240 3280 BY SIMILARITY,
FT REPEAT 3280 3320 BY SIMILARITY,
FT REPEAT 3320 3360 BY SIMILARITY,
FT REPEAT 3360 3400 BY SIMILARITY,
FT REPEAT 3400 3440 BY SIMILARITY,
FT REPEAT 3440 3480 BY SIMILARITY,
FT REPEAT 3480 3520 BY SIMILARITY,
FT REPEAT 3520 3560 BY SIMILARITY,
FT REPEAT 3560 3600 BY SIMILARITY,
FT REPEAT 3600 3640 BY SIMILARITY,
FT REPEAT 3640 3680 BY SIMILARITY,
FT REPEAT 3680 3720 BY SIMILARITY,
FT REPEAT 3720 3760 BY SIMILARITY,
FT REPEAT 3760 3800 BY SIMILARITY,
FT REPEAT 3800 3840 BY SIMILARITY,
FT REPEAT 3840 3880 BY SIMILARITY,
FT REPEAT 3880 3920 BY SIMILARITY,
FT REPEAT 3920 3960 BY SIMILARITY,
FT REPEAT 3960 4000 BY SIMILARITY,
FT REPEAT 4000 4040 BY SIMILARITY,
FT REPEAT 4040 4080 BY SIMILARITY,
FT REPEAT 4080 4120 BY SIMILARITY,
FT REPEAT 4120 4160 BY SIMILARITY,
FT REPEAT 4160 4200 BY SIMILARITY,
FT REPEAT 4200 4240 BY SIMILARITY,
FT REPEAT 4240 4280 BY SIMILARITY,
FT REPEAT 4280 4320 BY SIMILARITY,
FT REPEAT 4320 4360 BY SIMILARITY,
FT REPEAT 4360 4400 BY SIMILARITY,
FT REPEAT 4400 4440 BY SIMILARITY,
FT REPEAT 4440 4480 BY SIMILARITY,
FT REPEAT 4480 4520 BY SIMILARITY,
FT REPEAT 4520 4560 BY SIMILARITY,
FT REPEAT 4560 4600 BY SIMILARITY,
FT REPEAT 4600 4640 BY SIMILARITY,
FT REPEAT 4640 4680 BY SIMILARITY,
FT REPEAT 4680 4720 BY SIMILARITY,
FT REPEAT 4720 4760 BY SIMILARITY,
FT REPEAT 4760 4800 BY SIMILARITY,
FT REPEAT 4800 4840 BY SIMILARITY,
FT REPEAT 4840 4880 BY SIMILARITY,
FT REPEAT 4880 4920 BY SIMILARITY,
FT REPEAT 4920 4960 BY SIMILARITY,
FT REPEAT 4960 5000 BY SIMILARITY,
FT REPEAT 5000 5040 BY SIMILARITY,
FT REPEAT 5040 5080 BY SIMILARITY,
FT REPEAT 5080 5120 BY SIMILARITY,
FT REPEAT 5120 5160 BY SIMILARITY,
FT REPEAT 5160 5200 BY SIMILARITY,
FT REPEAT 5200 5240 BY SIMILARITY,
FT REPEAT 5240 5280 BY SIMILARITY,
FT REPEAT 5280 5320 BY SIMILARITY,
FT REPEAT 5320 5360 BY SIMILARITY,
FT REPEAT 5360 5400 BY SIMILARITY,
FT REPEAT 5400 5440 BY SIMILARITY,
FT REPEAT 5440 5480 BY SIMILARITY,
FT REPEAT 5480 5520 BY SIMILARITY,
FT REPEAT 5520 5560 BY SIMILARITY,
FT REPEAT 5560 5600 BY SIMILARITY,
FT REPEAT 5600 5640 BY SIMILARITY,
FT REPEAT 5640 5680 BY SIMILARITY,
FT REPEAT 5680 5720 BY SIMILARITY,
FT REPEAT 5720 5760 BY SIMILARITY,
FT REPEAT 5760 5800 BY SIMILARITY,
FT REPEAT 5800 5840 BY SIMILARITY,
FT REPEAT 5840 5880 BY SIMILARITY,
FT REPEAT 5880 5920 BY SIMILARITY,
FT REPEAT 5920 5960 BY SIMILARITY,
FT REPEAT 5960 6000 BY SIMILARITY,
FT REPEAT 6000 6040 BY SIMILARITY,
FT REPEAT 6040 6080 BY SIMILARITY,
FT REPEAT 6080 6120 BY SIMILARITY,
FT REPEAT 6120 6160 BY SIMILARITY,
FT REPEAT 6160 6200 BY SIMILARITY,
FT REPEAT 6200 6240 BY SIMILARITY,
FT REPEAT 6240 6280 BY SIMILARITY,
FT REPEAT 6280 6320 BY SIMILARITY,
FT REPEAT 6320 6360 BY SIMILARITY,
FT REPEAT 6360 6400 BY SIMILARITY,
FT REPEAT 6400 6440 BY SIMILARITY,
FT REPEAT 6440 6480 BY SIMILARITY,
FT REPEAT 6480 6520 BY SIMILARITY,
FT REPEAT 6520 6560 BY SIMILARITY,
FT REPEAT 6560 6600 BY SIMILARITY,
FT REPEAT 6600 6640 BY SIMILARITY,
FT REPEAT 6640 6680 BY SIMILARITY,
FT REPEAT 6680 6720 BY SIMILARITY,
FT REPEAT 6720 6760 BY SIMILARITY,
FT REPEAT 6760 6800 BY SIMILARITY,
FT REPEAT 6800 6840 BY SIMILARITY,
FT REPEAT 6840 6880 BY SIMILARITY,
FT REPEAT 6880 6920 BY SIMILARITY,
FT REPEAT 6920 6960 BY SIMILARITY,
FT REPEAT 6960 7000 BY SIMILARITY,
FT REPEAT 7000 7040 BY SIMILARITY,
FT REPEAT 7040 7080 BY SIMILARITY,
FT REPEAT 7080 7120 BY SIMILARITY,
FT REPEAT 7120 7160 BY SIMILARITY,
FT REPEAT 7160 7200 BY SIMILARITY,
FT REPEAT 7200 7240 BY SIMILARITY,
FT REPEAT 7240 7280 BY SIMILARITY,
FT REPEAT 7280 7320 BY SIMILARITY,
FT REPEAT 7320 7360 BY SIMILARITY,
FT REPEAT 7360 7400 BY SIMILARITY,
FT REPEAT 7400 7440 BY SIMILARITY,
FT REPEAT 7440 7480 BY SIMILARITY,
FT REPEAT 7480 7520 BY SIMILARITY,
FT REPEAT 7520 7560 BY SIMILARITY,
FT REPEAT 7560 7600 BY SIMILARITY,
FT REPEAT 7600 7640 BY SIMILARITY,
FT REPEAT 7640 7680 BY SIMILARITY,
FT REPEAT 7680 7720 BY SIMILARITY,
FT REPEAT 7720 7760 BY SIMILARITY,
FT REPEAT 7760 7800 BY SIMILARITY,
FT REPEAT 7800 7840 BY SIMILARITY,
FT REPEAT 7840 7880 BY SIMILARITY,
FT REPEAT 7880 7920 BY SIMILARITY,
FT REPEAT 7920 7960 BY SIMILARITY,
FT REPEAT 7960 8000 BY SIMILARITY,
FT REPEAT 8000 8040 BY SIMILARITY,
FT REPEAT 8040 8080 BY SIMILARITY,
FT REPEAT 8080 8120 BY SIMILARITY,
FT REPEAT 8120 8160 BY SIMILARITY,
FT REPEAT 8160 8200 BY SIMILARITY,
FT REPEAT 8200 8240 BY SIMILARITY,
FT REPEAT 8240 8280 BY SIMILARITY,
FT REPEAT 8280 8320 BY SIMILARITY,
FT REPEAT 8320 8360 BY SIMILARITY,
FT REPEAT 8360 8400 BY SIMILARITY,
FT REPEAT 8400 8440 BY SIMILARITY,
FT REPEAT 8440 8480 BY SIMILARITY,
FT REPEAT 8480 8520 BY SIMILARITY,
FT REPEAT 8520 8560 BY SIMILARITY,
FT REPEAT 8560 8600 BY SIMILARITY,
FT REPEAT 8600 8640 BY SIMILARITY,
FT REPEAT 8640 8680 BY SIMILARITY,
FT REPEAT 8680 8720 BY SIMILARITY,
FT REPEAT 8720 8760 BY SIMILARITY,
FT REPEAT 8760 8800 BY SIMILARITY,
FT REPEAT 8800 8840 BY SIMILARITY,
FT REPEAT 8840 8880 BY SIMILARITY,
FT REPEAT 8880 8920 BY SIMILARITY,
FT REPEAT 8920 8960 BY SIMILARITY,
FT REPEAT 8960 9000 BY SIMILARITY,
FT REPEAT 9000 9040 BY SIMILARITY,
FT REPEAT 9040 9080 BY SIMILARITY,
FT REPEAT 9080 9120 BY SIMILARITY,
FT REPEAT 9120 9160 BY SIMILARITY,
FT REPEAT 9160 9200 BY SIMILARITY,
FT REPEAT 9200 9240 BY SIMILARITY,
FT REPEAT 9240 9280 BY SIMILARITY,
FT REPEAT 9280 9320 BY SIMILARITY,
FT REPEAT 9320 9360 BY SIMILARITY,
FT REPEAT 9360 9400 BY SIMILARITY,
FT REPEAT 9400 9440 BY SIMILARITY,
FT REPEAT 9440 9480 BY SIMILARITY,
FT REPEAT 9480 9520 BY SIMILARITY,
FT REPEAT 9520 9560 BY SIMILARITY,
FT REPEAT 9560 9600 BY SIMILARITY,
FT REPEAT 9600 9640 BY SIMILARITY,
FT REPEAT 9640 9680 BY SIMILARITY,
FT REPEAT 9680 9720 BY SIMILARITY,
FT REPEAT 9720 9760 BY SIMILARITY,
FT REPEAT 9760 9800 BY SIMILARITY,
FT REPEAT 9800 9840 BY SIMILARITY,
FT REPEAT 9840 9880 BY SIMILARITY,
FT REPEAT 9880 9920 BY SIMILARITY,
FT REPEAT 9920 9960 BY SIMILARITY,
FT REPEAT 9960 10000 BY SIMILARITY,
```

```
FT DISULFID 439 448 BY SIMILARITY,
FT DISULFID 455 466 BY SIMILARITY,
FT DISULFID 460 475 BY SIMILARITY,
FT DISULFID 477 486 BY SIMILARITY,
FT DISULFID 493 504 BY SIMILARITY,
FT DISULFID 498 513 BY SIMILARITY,
FT DISULFID 515 524 BY SIMILARITY,
FT DISULFID 531 542 BY SIMILARITY,
FT DISULFID 536 551 BY SIMILARITY,
FT DISULFID 553 562 BY SIMILARITY,
FT DISULFID 569 579 BY SIMILARITY,
FT DISULFID 574 588 BY SIMILARITY,
FT DISULFID 590 599 BY SIMILARITY,
FT DISULFID 606 617 BY SIMILARITY,
FT DISULFID 611 626 BY SIMILARITY,
FT DISULFID 628 637 BY SIMILARITY,
FT DISULFID 644 654 BY SIMILARITY,
FT DISULFID 649 663 BY SIMILARITY,
FT DISULFID 665 674 BY SIMILARITY,
FT DISULFID 681 692 BY SIMILARITY,
FT DISULFID 686 701 BY SIMILARITY,
FT DISULFID 703 712 BY SIMILARITY,
FT DISULFID 719 729 BY SIMILARITY,
FT DISULFID 724 738 BY SIMILARITY,
FT DISULFID 740 749 BY SIMILARITY,
FT DISULFID 756 767 BY SIMILARITY,
FT DISULFID 761 776 BY SIMILARITY,
FT DISULFID 778 787 BY SIMILARITY,
FT DISULFID 794 805 BY SIMILARITY,
FT DISULFID 799 814 BY SIMILARITY,
FT DISULFID 816 825 BY SIMILARITY,
FT DISULFID 832 843 BY SIMILARITY,
FT DISULFID 837 854 BY SIMILARITY,
FT DISULFID 856 865 BY SIMILARITY,
FT DISULFID 872 883 BY SIMILARITY,
FT DISULFID 877 892 BY SIMILARITY,
FT DISULFID 894 903 BY SIMILARITY,
FT DISULFID 910 921 BY SIMILARITY,
FT DISULFID 915 930 BY SIMILARITY,
FT DISULFID 932 941 BY SIMILARITY,
FT DISULFID 966 997 BY SIMILARITY,
FT DISULFID 991 1006 BY SIMILARITY,
FT DISULFID 1008 1017 BY SIMILARITY,
FT DISULFID 1024 1035 BY SIMILARITY,
FT DISULFID 1029 1044 BY SIMILARITY,
FT DISULFID 1046 1055 BY SIMILARITY,
FT DISULFID 1062 1073 BY SIMILARITY,
FT DISULFID 1067 1082 BY SIMILARITY,
FT DISULFID 1084 1093 BY SIMILARITY,
FT DISULFID 1100 1121 BY SIMILARITY,
FT DISULFID 1115 1130 BY SIMILARITY,
FT DISULFID 1132 1141 BY SIMILARITY,
FT DISULFID 1148 1159 BY SIMILARITY,
FT DISULFID 1153 1168 BY SIMILARITY,
FT DISULFID 1170 1179 BY SIMILARITY,
FT DISULFID 1186 1197 BY SIMILARITY,
FT DISULFID 1206 1217 BY SIMILARITY,
FT DISULFID 1217 1244 BY SIMILARITY,
FT DISULFID 1224 1253 BY SIMILARITY,
FT DISULFID 1237 1252 BY SIMILARITY,

Query Match 25.7%; Score 324.5; DB 1; Length 2524;
Best Local Similarity 32.3%; Pred. No. 3e-19; Indels 71; Gaps 6;
Matches 64; Conservative 14; Mismatches 49;

OY 24 DICPNPCNGGICISGLADDSFSCCEGFGAGPNCSSVVEVASEDEKPTSAAGPCIPNPC 83
DB 908 DDCPPNPCHNGSCSCSGI--NMFPNCMPAGFRGPCEEDI-----NECASNPC 953
OY 84 HNGGT-----CEISEARNGTFTIG----LYCKQPRGFN 112
DB 954 KNGANCTDCVNSYTCOPGFSGICHESNTPDCTESSCFNGGTCTIDGINTFTYCCPGRFT 1013
```



```

FT DISULFID 244 253 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY.
FT DISULFID 265 280 BY SIMILARITY.
FT DISULFID 282 291 BY SIMILARITY.
FT DISULFID 298 311 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY.
FT DISULFID 322 331 BY SIMILARITY.
FT DISULFID 338 349 BY SIMILARITY.
FT DISULFID 343 358 BY SIMILARITY.
FT DISULFID 360 369 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY.
FT DISULFID 399 408 BY SIMILARITY.
FT DISULFID 415 428 BY SIMILARITY.
FT DISULFID 422 437 BY SIMILARITY.
FT DISULFID 439 448 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 460 475 BY SIMILARITY.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 503 BY SIMILARITY.
FT DISULFID 498 512 BY SIMILARITY.
FT DISULFID 514 523 BY SIMILARITY.
FT DISULFID 530 541 BY SIMILARITY.
FT DISULFID 535 550 BY SIMILARITY.
FT DISULFID 552 561 BY SIMILARITY.
FT DISULFID 568 578 BY SIMILARITY.
FT DISULFID 573 587 BY SIMILARITY.
FT DISULFID 589 598 BY SIMILARITY.
FT DISULFID 605 616 BY SIMILARITY.
FT DISULFID 610 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 653 BY SIMILARITY.
FT DISULFID 648 662 BY SIMILARITY.
FT DISULFID 664 673 BY SIMILARITY.
FT DISULFID 680 691 BY SIMILARITY.
FT DISULFID 685 700 BY SIMILARITY.
FT DISULFID 702 711 BY SIMILARITY.
FT DISULFID 718 728 BY SIMILARITY.
FT DISULFID 723 737 BY SIMILARITY.
FT DISULFID 739 748 BY SIMILARITY.
FT DISULFID 755 766 BY SIMILARITY.
FT DISULFID 760 775 BY SIMILARITY.
FT DISULFID 773 786 BY SIMILARITY.
FT DISULFID 793 804 BY SIMILARITY.
FT DISULFID 798 813 BY SIMILARITY.
FT DISULFID 815 824 BY SIMILARITY.
FT DISULFID 831 842 BY SIMILARITY.
FT DISULFID 836 853 BY SIMILARITY.
FT DISULFID 855 864 BY SIMILARITY.
FT DISULFID 871 882 BY SIMILARITY.
FT DISULFID 876 891 BY SIMILARITY.
FT DISULFID 893 902 BY SIMILARITY.
FT DISULFID 909 920 BY SIMILARITY.
FT DISULFID 914 929 BY SIMILARITY.
FT DISULFID 931 940 BY SIMILARITY.
FT DISULFID 947 958 BY SIMILARITY.
FT DISULFID 962 978 BY SIMILARITY.
FT DISULFID 969 978 BY SIMILARITY.
FT DISULFID 1023 1034 BY SIMILARITY.
FT DISULFID 1028 1043 BY SIMILARITY.
FT DISULFID 1045 1054 BY SIMILARITY.
FT DISULFID 1061 1072 BY SIMILARITY.
FT DISULFID 1066 1081 BY SIMILARITY.
FT DISULFID 1083 1092 BY SIMILARITY.
FT DISULFID 1099 1120 BY SIMILARITY.

```

Query Match 25.4%; Score 321; Length 2437;

Best Local Similarity 35.8%; Pred. No. 5, 6e-19;

Matches 62; Conservative 16; Mismatches 55; Indels 40; Gaps 4;

QY 26 CRRNPGNGICISGLADSFSCCEBGFAGPNCSSVEVASSEKPTSNAG----- 76
 DB 831 CSBRPCKNGVCHRESDFOSFCNCPAGMOGOTCE--VDINECVRNPCITNGVCENTLRGG 888

```

QY 77 -----PCIRPCHNGTCTEISEAYRDTPTIGYCKCPRGNGIHQ 117
DB 889 FQCRNPGFTGALCENDIDDCFEPNCSNGGVCO-----DRVNGFVCYCLAFRBERCA 941
QY 118 HNIIECAEPCRRNGICTDLVAVNSCECPGEFGRNOYK---CSGHLTEGG 167
DB 942 EDIECVASAPCRNGNCTDCVNSTYCSCPAFSGINCEINTEPTDCTESSCPNGG 994

RESULT 7
FBP1_STRPU STANDARD; PRT; 1064 AA.
ID FBP1_STRPU
AC P10079;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FIBROPELIN I PRECURSOR (EPIDERMAL GROWTH FACTOR-RELATED PROTEIN 1)
DE (UEGF-1).
GN EGF1.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinzoa;
OC Echinoidea; Euechinoidea; Echinacea; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90112459; PubMed=2514273;
RA Delgadillo-Reynoso M.G., Rollo D.R., Hursh D.A., Raff R.A.;
RT "Structural analysis of the uegf gene in the sea urchin
strongylocentrotus purpuratus reveals more similarity to vertebrate
than to invertebrate genes with EGF-like repeats."
RL J. Mol. Evol. 29:314-327(1989).
RN [2]
RP SEQUENCE OF 279-476 AND 781-1064 FROM N.A.
RX MEDLINE=87319677; PubMed=3498216;
RA Hursh D.A., Andrews M.E., Raff R.A.;
RT "A sea urchin gene encodes a polypeptide homologous to epidermal
growth factor."
RL Science 237:1487-1490(1987).
RN [3]
RP AVIDIN-LIKE DOMAIN.
RX MEDLINE=89196806; PubMed=2784773;
RA Hunt L.T., Barker W.C.;
RT "Avidin-like domain in an epidermal growth factor homolog from a sea
urchin."
RL FASEB J. 3:1760-1764(1989).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=91285254; PubMed=2060714;
RA Bisgrove B.W., Andrews M.E., Raff R.A.;
RT "Fibropellins, products of an EGF repeat-containing gene, form a
unique extracellular matrix structure that surrounds the sea urchin
embryo."
RL Dev. Biol. 146:89-99(1991).
RN [5]
RP FUNCTION: FORM THE APICAL LAMINA, A COMPONENT OF THE EXTRACELLULAR
MATRIX.
CC -I- SUBCELLULAR LOCATION: EXTRACELLULAR, IN VESICLES IN THE CYTOPLASM
OR UNFERTILIZED EGGS, THEN TO THE BASE OF THE HYALIN LAYER
THROUGHOUT DEVELOPMENT AND FINALLY IN THE APICAL LAMINA IN LATE
EMBRYOS AND EARLY LARVAE.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
PRODUCED BY ALTERNATIVE SPLICING. THE SMALL FORM (1B) LACKS 8 EGF
REPEATS.
CC -I- DEVELOPMENTAL STAGE: MODERATE LEVELS IN UNFERTILIZED EGGS AND
DURING EARLY CLEAVAGE, THEN RAPIDLY INCREASES IN ABUNDANCE BETWEEN
LATE MORULA AND MESENCEPHALIC BLASTULA STAGES TO MAXIMAL LEVELS
MAINTAINED THROUGH SUBSEQUENT STAGES. EXPRESSED BOTH MATERNALLY
AND ZYGOTICALLY.
CC -I- SIMILARITY: CONTAINS 21 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS 1 CUB DOMAIN.
CC -I- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR
TO AVIDIN/STREPTAVIDIN.

```

Query Match	Best Local Similarity	Score	DB 1	Length	1064
Matches 70; Conservative	34.7%	25.2%	35.7%	16;	51; Gaps
OY 7	AWLVLGSLGPOFG-----KGDICNENPCNGGICLSGLADSFSCCEPGFAGPNC	59			
DB 154	AMWFFSTRIVNRGRITFFSSDGDCCDPCNLQNGAAC-TDLVND-YACTCPCPGTGRNC	211			
OY 60	SSVEVAASDEKRTSGACPIPNPCINGGICELSEAYRBDFTITGVCKCPGFGNGIHCRHN	119			
DB 212	-----EIDIDE-----CASDPCQNGGACV-----DGVNGVCMCVGFDGDECCNN	252			
OY 120	INBOEAPFNGGICTDLVANSCECPGFGPNC-----	156			
DB 253	INBOEAPFNGGICTDLVANSCECPGFGPNC-----	156			
OY 253	INBOEAPFNGGICTDLVANSCECPGFGPNC-----	156			

OY 157 KCSGHLGIEGIIISNOQITASS 178
 DB 313 TCSGHLGIEGIIISNOQITASS 334

RESULT 8
 FBP3_STRPU STANDARD; PRT; 570 AA.
 ID FBP3_STRPU
 AC P49013;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE FIBROPELIN C PRECURSOR (EPIDERMAL GROWTH FACTOR-RELATED PROTEIN 3)
 DE (EGF III) (FIBROPELIN III).
 GN EGF3.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxId=7668;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-gastrula;
 RA MEDLINE=93273088; PubMed=8500658;
 RA Bisgrove B.W., Rafi R.A.;
 RT "The SPECf III gene encodes a member of the fibropeilins: EGF repeat-
 containing proteins that form the apical lamina of the sea urchin
 embryo";
 RL Dev. Biol. 157:526-538(1993).
 CC -1- FUNCTION: FORM THE APICAL LAMINA, A COMPONENT OF THE EXTRACELLULAR
 MATRIX.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN UNFERTILIZED EGGS AND DURING
 MORULA AND MESENCHYME BLASTULA STAGES TO MAXIMAL LEVELS MAINTAINED
 THROUGH SUBSEQUENT STAGES.
 CC -1- MISCELLANEOUS: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR
 TO AVIDIN/STREPTAVIDIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L07045; AAA30045.1; -
 DR HSP: P00740; 11XA.
 DR InterPro: IPR000088; -
 DR InterPro: IPR000152; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR000859; -
 DR InterPro: IPR001438; -
 DR InterPro: IPR001881; -
 DR Pfam: PF01382; Avidin; 1.
 DR Pfam: PF00431; CUB; 1.
 DR Pfam: PF00068; EGF; 8.
 DR PRINTS: PR00010; EGFBL00D.
 DR PROSITE: PS00010; ASX_HYDROXYL; 8.
 DR PROSITE: PS00022; EGF_1; 8.
 DR PROSITE: PS00577; AVIDIN; 1.
 DR PROSITE: PS01180; CUB; 1.
 DR PROSITE: PS01186; EGF_2; 7.
 DR PROSITE: PS01187; EGF_CA; 6.
 DR Bioclin; EGF-like domain; Repeat; Signal; Glycoprotein;
 KW Calcium-binding.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 570 FIBROPELIN C.
 FT DOMAIN 18 55 EGF-LIKE 1.

FT DOMAIN 62 175 CUB.
 FT DOMAIN 176 212 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 214 250 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 252 288 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 290 326 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 328 364 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 366 402 EGF-LIKE 7.
 FT DOMAIN 404 440 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 442 570 AVIDIN-LIKE.
 FT DISULFID 23 34 BY SIMILARITY.
 FT DISULFID 28 43 BY SIMILARITY.
 FT DISULFID 45 54 BY SIMILARITY.
 FT DISULFID 180 191 BY SIMILARITY.
 FT DISULFID 185 200 BY SIMILARITY.
 FT DISULFID 202 211 BY SIMILARITY.
 FT DISULFID 218 229 BY SIMILARITY.
 FT DISULFID 223 238 BY SIMILARITY.
 FT DISULFID 240 249 BY SIMILARITY.
 FT DISULFID 256 267 BY SIMILARITY.
 FT DISULFID 261 276 BY SIMILARITY.
 FT DISULFID 278 287 BY SIMILARITY.
 FT DISULFID 294 305 BY SIMILARITY.
 FT DISULFID 299 314 BY SIMILARITY.
 FT DISULFID 316 325 BY SIMILARITY.
 FT DISULFID 332 343 BY SIMILARITY.
 FT DISULFID 337 352 BY SIMILARITY.
 FT DISULFID 354 363 BY SIMILARITY.
 FT DISULFID 370 381 BY SIMILARITY.
 FT DISULFID 375 390 BY SIMILARITY.
 FT DISULFID 392 401 BY SIMILARITY.
 FT DISULFID 408 419 BY SIMILARITY.
 FT DISULFID 413 428 BY SIMILARITY.
 FT DISULFID 430 439 BY SIMILARITY.
 FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 570 AA: 6116 MW: 86553E1C05E6E6 CRC64;
 Query Match 24.8%; Score 313.5; DB 1; Length 570;
 Best Local Similarity 30.6%; Pred. No. 6.76-19;
 Matches 64; Conservative 18; Mismatches 58; Indels 69; Gaps 6;
 OY 7 ANLVGLSLGVQFG-----KEDICNPCENGICLSGLADSFCECEGPGAGNC 59
 DB 154 AWFISTDNRINRKRFRITSSDGDCTPNCPLNGATVDVND--YQCTAPGFTGNC 211
 OY 60 SSVVEVASDEEKPTSGPCIPNPNCHNGTCEISEAYRSDTFIGYCKPGRFGNIHQHN 119
 DB 212 E-----TDIDECASAPCRNGACV-----DQVNGYTCNCLPGNVCENN 252
 OY 120 -----INCEAEPCCRNIGICTDVLVNY 141
 DB 253 INECASIPCLNGICVDGINPFACTCLPGYTGILCETDINECASSPCONGSCTDAVNY 312
 OY 142 SCEGPEFMGRNQ---YKCSGHLGIEG 167
 DB 313 TDCCRAGFTGNCNCTININECASSPCLNGG 341
 RESULT 9
 NCBI_MOUSE
 ID NCBI_MOUSE STANDARD; PRT; 2531 AA.
 AC Q01705;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (NOTCH PROTEIN).
 GN NOTCH1 OR NOTCH.
 OS Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Embryo:
RA MEDLINE-93194170: PubMed-8449489:
RA Franco del Amo F., Gendron-Maguire M., Swiatek P.J., Jenkins N.A.,
RA Copeland N.G., Gridley T.;
RT "Cloning, analysis, and chromosomal localization of Notch-1, a mouse
RT homolog of Drosophila Notch."
RL Genomics 15:259-264(1993).
[2]
RN RP SEQUENCE OF 1551-2170 FROM N.A.
RC TISSUE-Embryo:
RA MEDLINE-93048835: PubMed-1425352:
RA Franco del Amo F., Smith D.E., Swiatek P.J., Gendron-Maguire M.,
RA Greenspan R.J., McMahon A.P., Gridley T.;
RT "Expression pattern of Notch, a mouse homolog of Drosophila Notch,
RT suggests an important role in early postimplantation mouse
RT development."
RL Development 115:737-744(1992).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z11886; CAAT7941.1; -
DR HSPB: P00740; ITXA.
DR MCD: MG1:97363; Notch1.
DR InterPro: IPR000152; -
DR InterPro: IPR000561; -
DR InterPro: IPR000800; -
DR InterPro: IPR001438; -
DR InterPro: IPR001881; -
DR InterPro: IPR002110; -
DR Pfam: PF00008; EGF; 35.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00010; EGFBL00D.
DR PROSITE: PSS0088; ANK_REPEAT; 2.
DR PROSITE: PSS0297; ANK_REPEAT_REGION; 1.
DR PROSITE: PSS0022; ASX_HYDROXYL; 22.
DR PROSITE: PSS0186; EGF_1; 34.
DR PROSITE: PSS0187; EGF_2; 27.
DR PROSITE: PSS0187; EGF_CA; 21.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 18
FT CHAIN 1 2531
FT FT 19 1725 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
FT DOMAIN 1726 1746 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1747 2531 POTENTIAL.
FT DOMAIN 1747 2531 CYTOPLASMIC (POTENTIAL).
FT FT 20 58
FT DOMAIN 20 59 EGF-LIKE 1.
FT FT 102 139 EGF-LIKE 2.
FT DOMAIN 102 139 EGF-LIKE 3.
FT FT 140 176 EGF-LIKE 4.
FT DOMAIN 140 176 EGF-LIKE 5.
FT FT 218 255 EGF-LIKE 6.
FT DOMAIN 218 255 EGF-LIKE 7.
FT FT 257 293 EGF-LIKE 8.
FT DOMAIN 257 293 EGF-LIKE 9.
FT FT 333 371 EGF-LIKE 10.
FT DOMAIN 333 371 EGF-LIKE 11.
FT FT 372 410 EGF-LIKE 12.
FT DOMAIN 372 410 EGF-LIKE 13.
FT FT 412 450 EGF-LIKE 14.
FT DOMAIN 412 450 EGF-LIKE 15.
FT FT 452 488 EGF-LIKE 16.
FT DOMAIN 452 488 EGF-LIKE 17.
FT FT 490 526 EGF-LIKE 18.

```

```

FT DOMAIN 528 564 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 566 601 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 603 639 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 641 676 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 678 714 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 716 751 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 753 789 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 791 827 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 829 867 EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 869 905 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 907 943 EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 945 981 EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 983 1019 EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1021 1057 EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1059 1095 EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1097 1143 EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1145 1181 EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1183 1219 EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1221 1265 EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1267 1305 EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1307 1346 EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1348 1384 EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1387 1426 EGF-LIKE 36, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1449 1462 CYS-RICH.
FT REPEAT 1449 1462 CYS-RICH.
FT REPEAT 1481 1522 LIN/NOTCH 1.
FT REPEAT 1523 1562 LIN/NOTCH 2.
FT REPEAT 1562 1597 LIN/NOTCH 3.
FT REPEAT 1597 1647 ANK 1.
FT REPEAT 1647 1699 ANK 2.
FT REPEAT 1699 1751 ANK 3.
FT REPEAT 1751 1803 ANK 4.
FT REPEAT 1803 1855 ANK 5.
FT DISULFID 24 37 BY SIMILARITY.
FT DISULFID 31 46 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 68 87 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 111 127 BY SIMILARITY.
FT DISULFID 129 138 BY SIMILARITY.
FT DISULFID 144 155 BY SIMILARITY.
FT DISULFID 149 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 182 195 BY SIMILARITY.
FT DISULFID 189 204 BY SIMILARITY.
FT DISULFID 206 215 BY SIMILARITY.
FT DISULFID 222 233 BY SIMILARITY.
FT DISULFID 227 243 BY SIMILARITY.
FT DISULFID 245 254 BY SIMILARITY.
FT DISULFID 261 272 BY SIMILARITY.
FT DISULFID 266 281 BY SIMILARITY.
FT DISULFID 283 292 BY SIMILARITY.
FT DISULFID 299 312 BY SIMILARITY.
FT DISULFID 312 321 BY SIMILARITY.
FT DISULFID 323 332 BY SIMILARITY.
FT DISULFID 339 350 BY SIMILARITY.
FT DISULFID 344 359 BY SIMILARITY.
FT DISULFID 361 370 BY SIMILARITY.
FT DISULFID 376 387 BY SIMILARITY.
FT DISULFID 381 398 BY SIMILARITY.
FT DISULFID 400 409 BY SIMILARITY.
FT DISULFID 416 429 BY SIMILARITY.
FT DISULFID 423 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 446 456 BY SIMILARITY.
FT DISULFID 461 476 BY SIMILARITY.
FT DISULFID 478 487 BY SIMILARITY.
FT DISULFID 494 505 BY SIMILARITY.
FT DISULFID 499 514 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 532 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 554 563 BY SIMILARITY.

```


FT	DISULFID	570	580		BY SIMILARITY.
FT	DISULFID	575	589		BY SIMILARITY.
FT	DISULFID	591	600		BY SIMILARITY.
FT	DISULFID	607	618		BY SIMILARITY.
FT	DISULFID	612	627		BY SIMILARITY.
FT	DISULFID	629	635		BY SIMILARITY.
FT	DISULFID	645	658		BY SIMILARITY.
FT	DISULFID	650	664		BY SIMILARITY.
FT	DISULFID	666	675		BY SIMILARITY.
FT	DISULFID	682	693		BY SIMILARITY.
FT	DISULFID	687	702		BY SIMILARITY.
FT	DISULFID	704	713		BY SIMILARITY.
FT	DISULFID	720	730		BY SIMILARITY.
FT	DISULFID	725	739		BY SIMILARITY.
FT	DISULFID	741	750		BY SIMILARITY.
FT	DISULFID	757	768		BY SIMILARITY.
FT	DISULFID	762	777		BY SIMILARITY.
FT	DISULFID	779	788		BY SIMILARITY.
FT	DISULFID	795	806		BY SIMILARITY.
FT	DISULFID	800	815		BY SIMILARITY.
FT	DISULFID	817	826		BY SIMILARITY.
FT	DISULFID	833	844		BY SIMILARITY.
FT	DISULFID	838	855		BY SIMILARITY.
FT	DISULFID	857	866		BY SIMILARITY.
FT	DISULFID	873	884		BY SIMILARITY.
FT	DISULFID	878	893		BY SIMILARITY.
FT	DISULFID	895	904		BY SIMILARITY.
FT	DISULFID	911	922		BY SIMILARITY.
FT	DISULFID	916	931		BY SIMILARITY.
FT	DISULFID	933	942		BY SIMILARITY.
FT	DISULFID	987	998		BY SIMILARITY.
FT	DISULFID	992	1007		BY SIMILARITY.
FT	DISULFID	1009	1018		BY SIMILARITY.
FT	DISULFID	1025	1036		BY SIMILARITY.
FT	DISULFID	1030	1045		BY SIMILARITY.
FT	DISULFID	1047	1056		BY SIMILARITY.
FT	DISULFID	1063	1074		BY SIMILARITY.
FT	DISULFID	1068	1083		BY SIMILARITY.
FT	DISULFID	1085	1094		BY SIMILARITY.
FT	DISULFID	1101	1122		BY SIMILARITY.
FT	DISULFID	1116	1131		BY SIMILARITY.
FT	DISULFID	1133	1142		BY SIMILARITY.
FT	DISULFID	1149	1160		BY SIMILARITY.

Query Match Best Local Similarity 24.8%; Score 312.5; DB 1; Length 2531;
Matches 63; Conservative 12; Mismatches 42; Indels 55; Gaps 5

QY	23	GDIICNP-----PENGIGISGLADDSFSCDEPFAFPNCSVVEASDEKRTSA	75
Db	672	GSMCNVNIDEACGSPCHNGTCEGDA--GFYCRCPDEGYHDPCLT-----EVCNCS	722
QY	76	GPCI-----PPRCHNGGCETSEAYRDPRFY	103
Db	723	NPCIHACBRDLNGKYRCDCAPGMSTNCDDINNECESNPCVNGGTCK-----DMSGT	775
QY	104	VCKRCPEFGSINGHCQHNNINECEAEPCRNAGICTDLYAVNTSCBEPGEFMGRNQ	155
Db	776	VCMCREGFSGPCNQTGINECASNPCINCNOOTCTIDVAGKCYCPLPYTGANCE	827

RESULT 10
NTCI_RAT NTCL_RAT STANDARD; PRT; 2531 AA.
AC 007008;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR.
GN NOTCH1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

RN NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Schwann cell.;
RX MEDLINE=92111383; PubMed=1764995;
RA Weinmaster G., Roberts V.J., Lemke G.;
RT "A homolog of Drosophila Notch expressed during mammalian
RL development.";
RN Development 113:199-205(1991).
[2]
RN REVISIONS TO 1652-1653.
RP Weinmaster G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC -! FUNCTION. REQUIRED FOR THE CORRECT DIFFERENTIATION OF A NUMBER
OF TISSUES.
CC -! SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -! DEVELOPMENTAL STAGE: IN THE EMBryo, HIGHEST LEVELS OCCUR BETWEEN
DAYS 12 AND 14 AND DECREASE RAPIDLY TO MUCH LOWER LEVELS IN THE
ADULT.
CC -! SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC -! SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC -! SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -! SIMILARITY: CONTAINS 5 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).

DR EMBL; X57405; CAA40667.1; -.
DR HSSP; P00740; 1IXA.
DR InterPro; IPRO00152; .
DR InterPro; IPRO000561; .
DR InterPro; IPRO000800; .
DR InterPro; IPRO01438; .
DR InterPro; IPRO01881; .
DR InterPro; IPRO02049; .
DR InterPro; IPRO02110; .
DR Pfam; PF00008; EGF; 36.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBLOOD.
DR PRINTS; PR00011; EGFLAMININ.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; 35.
DR PROSITE; PS01186; EGF_2; 26.
DR PROSITE; PS01187; EGF_CA; 21.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
Transmembrane; Signal; Glycoprotein.

KW SIGNAL.
FT CHAIN 1 .. 18 POTENTIAL.
FT CHAIN 19 .. 2531 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
FT DOMAIN 19 .. 1723 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1724 .. 1746 POTENTIAL.
FT DOMAIN 1747 .. 2531 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 .. 58
FT DOMAIN 59 .. 99 EGF-Like 1.
FT DOMAIN 102 .. 139 EGF-Like 2.
FT DOMAIN 140 .. 176 EGF-Like 3.
FT DOMAIN 178 .. 216 EGF-Like 4.
FT DOMAIN 218 .. 255 EGF-Like 5.
FT DOMAIN 257 .. 293 EGF-Like 6.
FT DOMAIN 295 .. 333 EGF-Like 7.
FT DOMAIN 335 .. 371 EGF-Like 8.
FT DOMAIN 372 .. 410 EGF-Like 9.
FT DOMAIN 412 .. 450 EGF-Like 10.
FT DOMAIN 452 .. 488 CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 490 .. 526 EGF-Like 11.
FT DOMAIN 528 .. 564 EGF-Like 12.
EGF-Like 13, CALCIUM-BINDING (POTENTIAL).
EGF-Like 14, CALCIUM-BINDING (POTENTIAL).
EGF-Like 1, CALCIUM-BINDING (POTENTIAL).
EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
EGF-Like 5, CALCIUM-BINDING (POTENTIAL).
EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
EGF-Like 7, CALCIUM-BINDING (POTENTIAL).
EGF-Like 8, CALCIUM-BINDING (POTENTIAL).
EGF-Like 9, CALCIUM-BINDING (POTENTIAL).
EGF-Like 10, CALCIUM-BINDING (POTENTIAL).
EGF-Like 11, CALCIUM-BINDING (POTENTIAL).
EGF-Like 12, CALCIUM-BINDING (POTENTIAL).
EGF-Like 13, CALCIUM-BINDING (POTENTIAL).
EGF-Like 14, CALCIUM-BINDING (POTENTIAL).

```
FT DOMAIN 566 601 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 603 639 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 641 676 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 678 714 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 716 751 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 753 789 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 791 827 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 829 867 EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 869 905 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 907 943 EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 945 981 EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 983 1019 EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1021 1057 EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1059 1095 EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1097 1143 EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1145 1181 EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1183 1219 EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1221 1265 EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1267 1305 EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1307 1346 EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1348 1384 EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1387 1426 EGF-LIKE 36, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1449 1462 CTS-RICH.
FT REPEAT 1917 1946 ANK 1.
FT REPEAT 1950 1980 ANK 2.
FT REPEAT 1984 2013 ANK 3.
FT REPEAT 2017 2046 ANK 4.
FT REPEAT 2050 2079 ANK 5.
FT DISULFID 24 37 BY SIMILARITY.
FT DISULFID 31 46 BY SIMILARITY.
FT DISULFID 48 57 BY SIMILARITY.
FT DISULFID 63 74 BY SIMILARITY.
FT DISULFID 68 87 BY SIMILARITY.
FT DISULFID 89 98 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 111 127 BY SIMILARITY.
FT DISULFID 129 138 BY SIMILARITY.
FT DISULFID 144 155 BY SIMILARITY.
FT DISULFID 149 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 182 195 BY SIMILARITY.
FT DISULFID 189 204 BY SIMILARITY.
FT DISULFID 206 215 BY SIMILARITY.
FT DISULFID 222 233 BY SIMILARITY.
FT DISULFID 227 243 BY SIMILARITY.
FT DISULFID 245 254 BY SIMILARITY.
FT DISULFID 261 272 BY SIMILARITY.
FT DISULFID 266 281 BY SIMILARITY.
FT DISULFID 283 292 BY SIMILARITY.
FT DISULFID 299 312 BY SIMILARITY.
FT DISULFID 306 321 BY SIMILARITY.
FT DISULFID 323 332 BY SIMILARITY.
FT DISULFID 339 350 BY SIMILARITY.
FT DISULFID 344 359 BY SIMILARITY.
FT DISULFID 361 370 BY SIMILARITY.
FT DISULFID 376 387 BY SIMILARITY.
FT DISULFID 381 398 BY SIMILARITY.
FT DISULFID 400 409 BY SIMILARITY.
FT DISULFID 416 429 BY SIMILARITY.
FT DISULFID 423 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 456 467 BY SIMILARITY.
FT DISULFID 461 476 BY SIMILARITY.
FT DISULFID 478 487 BY SIMILARITY.
FT DISULFID 494 505 BY SIMILARITY.
FT DISULFID 499 514 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 532 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 554 563 BY SIMILARITY.
FT DISULFID 570 580 BY SIMILARITY.
FT DISULFID 575 589 BY SIMILARITY.
FT DISULFID 591 600 BY SIMILARITY.
```

```
FT DISULFID 607 618 BY SIMILARITY.
FT DISULFID 612 627 BY SIMILARITY.
FT DISULFID 629 638 BY SIMILARITY.
FT DISULFID 645 655 BY SIMILARITY.
FT DISULFID 650 664 BY SIMILARITY.
FT DISULFID 666 675 BY SIMILARITY.
FT DISULFID 682 693 BY SIMILARITY.
FT DISULFID 687 702 BY SIMILARITY.
FT DISULFID 704 713 BY SIMILARITY.
FT DISULFID 720 730 BY SIMILARITY.
FT DISULFID 725 739 BY SIMILARITY.
FT DISULFID 741 750 BY SIMILARITY.
FT DISULFID 757 768 BY SIMILARITY.
FT DISULFID 762 777 BY SIMILARITY.
FT DISULFID 779 788 BY SIMILARITY.
FT DISULFID 795 806 BY SIMILARITY.
FT DISULFID 800 815 BY SIMILARITY.
FT DISULFID 817 826 BY SIMILARITY.
FT DISULFID 833 844 BY SIMILARITY.
FT DISULFID 838 855 BY SIMILARITY.
FT DISULFID 857 866 BY SIMILARITY.
FT DISULFID 873 884 BY SIMILARITY.
FT DISULFID 878 893 BY SIMILARITY.
FT DISULFID 895 904 BY SIMILARITY.
FT DISULFID 911 922 BY SIMILARITY.
FT DISULFID 916 931 BY SIMILARITY.
FT DISULFID 933 942 BY SIMILARITY.
FT DISULFID 987 998 BY SIMILARITY.
FT DISULFID 992 1007 BY SIMILARITY.
FT DISULFID 1009 1018 BY SIMILARITY.
FT DISULFID 1025 1036 BY SIMILARITY.
FT DISULFID 1030 1045 BY SIMILARITY.
FT DISULFID 1047 1056 BY SIMILARITY.
FT DISULFID 1063 1074 BY SIMILARITY.
FT DISULFID 1068 1083 BY SIMILARITY.
FT DISULFID 1085 1094 BY SIMILARITY.
FT DISULFID 1101 1122 BY SIMILARITY.
FT DISULFID 1116 1131 BY SIMILARITY.
FT DISULFID 1133 1142 BY SIMILARITY.
FT DISULFID 1149 1160 BY SIMILARITY.
FT DISULFID 1154 1169 BY SIMILARITY.
FT DISULFID 1171 1180 BY SIMILARITY.
FT DISULFID 1187 1198 BY SIMILARITY.
FT DISULFID 1192 1207 BY SIMILARITY.

Query Match 24.8%; Score 312.5; DB 1; Length 2531;
Best Local Similarity 36.6%; Pred. No. 2.9e-18;
Matches 63; Conservative 12; Mismatches 42; Indels 55; Gaps 5;

QY 23 GDICNP-----PCENGICISGLADDSFSCPEGFAGPNCSSVEVASDEKPTSA 75
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 672 GSMCNVNIIDECAGSPHNGTCEGDGIA--GFTCRCEPGSHDPTCLS-----EVNECHS 722
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 76 GPCI-----PNCCHNGGTCEISEAYRGDTFTIGY 103
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 723 NPCIHACRDLNGYKCDAPGWSGTNCDIINNNECESNCPVNGTCK-----DWTSGY 775
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 104 VCKCPGPFNCHQHNNINCEAEPCRNNGTCTDVANYSCEPGEPMGNQC 155
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 776 VCTCRGEGFSGPNCQTNINCCASNPCLNGTCTIDVAGVGCNCPLEPYGTACE 827

RESULT 11
DLL1_RAT
ID DLL1_RAT STANDARD; PRT; 714 AA.
AC P97677;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE DELTA-LIKE PROTEIN 1 PRECURSOR (DROSOPHILA DELTA HOMOLOG 1) (DELTA).
GN DLL1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Disbio G., Hebsht L., Boulter J., Weinmaster G.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
 CC MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
 CC UNDERLYING SOMITOMYOTENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U78889; AAB37343.1. -
 DR HSP: P00740; 11XA.
 DR InterPro: IPR000152; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001438; -
 DR InterPro: IPR001774; -
 DR InterPro: IPR001881; -
 DR Pfam: PF01414; DSL; 1.
 DR Pfam: PF00008; EGF; 7.
 DR PRINTS: PRO0010; EGFBLD.
 DR PROSITE: PS00010; ASX_HYDROXYL; 3.
 DR PROSITE: PS00022; EGF_1; 8.
 DR PROSITE: PS01186; EGF-2; 8.
 DR PROSITE: PS01187; EGF_CA; 2.
 DR Signal: EGF-like domain; Repeat; Transmembrane; Developmental protein;
 KW Differentiation; Glycoprotein.
 FT SIGNAL 1 17
 FT CHAIN 18 714
 FT DOMAIN 18 537
 FT TRANSMEM 538 560
 FT DOMAIN 561 714
 FT DOMAIN 225 253
 FT DOMAIN 256 284
 FT DOMAIN 291 324
 FT DOMAIN 331 362
 FT DOMAIN 369 401
 FT DOMAIN 408 439
 FT DOMAIN 446 477
 FT DOMAIN 484 515
 FT DISULFID 225 236
 FT DISULFID 229 242
 FT DISULFID 244 253
 FT DISULFID 256 267
 FT DISULFID 262 273
 FT DISULFID 275 284
 FT DISULFID 291 303
 FT DISULFID 297 313
 FT DISULFID 315 324
 FT DISULFID 331 342
 FT DISULFID 336 351
 FT DISULFID 353 362
 FT DISULFID 369 380
 FT DISULFID 374 390
 FT DISULFID 392 401
 FT DISULFID 408 419
 FT DISULFID 413 428
 FT DISULFID 430 439
 FT DISULFID 446 466
 FT DISULFID 468 477
 FT DISULFID 484 495
 FT DISULFID 489 504

FT DISULFID 506 515 BY SIMILARITY.
 FT CARBOHYD 476 476 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 714 AA; 77378 MW; 4B8E2272BAEA27E CRC64;
 Query Match 24.4%; Score 307.5; DB 1; Length 714;
 Best local Similarity 29.5%; Pred. No. 2.5e-18;
 Matches 70; Conservative 26; Mismatches 80; Indels 61; Gaps 7;
 QY 24 DICNPNPCNGICGLADDSFSCPEGPAGPNCSSVVEVASDEBKPISAGCIPNPC 83
 DB 329 DECAPSPRCNGSGCTD--LEDSTCTCPPTGYKVC-----LSAMTCADGPC 374
 QY 84 HNGG-----TCEISEARYG-----DTFTGYCKCPRGF 111
 DB 375 FNGRCSDNPDGAYCHCPAGFSFNCERKIDLCSSPSCNGACVDLGNSTLCRCQTGF 434
 QY 112 NGIRHQHINCECAEPKNGGICIDLVANYSCEPGEFMGNCCQYKS--GHLDIEGII 169
 DB 435 SGRCYEDVDDCASSPCANGGTCRDSVNDPSCPTCPGYTGNCAPVSRCEHAPCHNG-- 492
 QY 170 SMOQTASSNHRALFGLQWVPYVAALNKKGLI-----NMTAENDRWPMIQTVG 221
 DB 493 -----ATCHQRCQRYMCECAQGYGANGCFLPEPPPDILVAAQGSFPMVAVCAG 543
 RESULT 12
 NOTC_DROME STANDARD: PRT: 2703 AA.
 AC P07207; P04154;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NEUROGENIC LOCUS NOTCH PROTEIN PRECURSOR.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86079539; PubMed=3935325;
 RA Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
 RT "Nucleotide sequence from the neurogenic locus notch implies a gene
 RT product that shares homology with proteins containing EGF-like
 RT repeats.";
 RL Cell 43:567-581(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OREGON-R;
 RX MEDLINE=87064624; PubMed=3097517;
 RA Kidd S., Kelley M.R., Young M.W.;
 RT "Sequence of the notch locus of Drosophila melanogaster: relationship
 RT of the encoded protein to mammalian clotting and growth factors.";
 RL Mol. Cell. Biol. 6:3094-3108(1986).
 RN [3]
 RP SEQUENCE OF 2505-2611 FROM N.A.
 RX MEDLINE=85093329; PubMed=2981631;
 RA Wharton K.A., Yelovonick B., Finnerty V.G., Artavanis-Tsakonas S.;
 RT "opa: a novel family of transcribed repeats shared by the Notch locus
 RT and other developmentally regulated loci in D. melanogaster.";
 RL Cell 40:55-62(1985).
 RN [4]
 RP SEQUENCE OF 1-8 FROM N.A.
 RX MEDLINE=87257846; PubMed=3037327;
 RA Kelley M.R., Kidd S., Berg R.L., Young M.W.;
 RT "Restriction of P-element insertions at the Notch locus of Drosophila
 RT melanogaster.";
 RL Mol. Cell. Biol. 7:1545-1548(1987).
 RN [5]
 RP REVIEW.
 RA Harris W.A.;


```
FT DISULFID 111 127 BY SIMILARITY.
FT DISULFID 129 138 BY SIMILARITY.
FT DISULFID 144 155 BY SIMILARITY.
FT DISULFID 149 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 182 195 BY SIMILARITY.
FT DISULFID 189 204 BY SIMILARITY.
FT DISULFID 206 215 BY SIMILARITY.
FT DISULFID 222 233 BY SIMILARITY.
FT DISULFID 227 243 BY SIMILARITY.
FT DISULFID 245 254 BY SIMILARITY.
FT DISULFID 261 272 BY SIMILARITY.
FT DISULFID 266 281 BY SIMILARITY.
FT DISULFID 283 292 BY SIMILARITY.
FT DISULFID 299 312 BY SIMILARITY.
FT DISULFID 306 321 BY SIMILARITY.
FT DISULFID 323 332 BY SIMILARITY.
FT DISULFID 339 350 BY SIMILARITY.
FT DISULFID 344 359 BY SIMILARITY.
FT DISULFID 361 370 BY SIMILARITY.
FT DISULFID 376 387 BY SIMILARITY.
FT DISULFID 381 398 BY SIMILARITY.
FT DISULFID 400 409 BY SIMILARITY.
FT DISULFID 416 429 BY SIMILARITY.
FT DISULFID 423 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 456 467 BY SIMILARITY.
FT DISULFID 461 476 BY SIMILARITY.
FT DISULFID 478 487 BY SIMILARITY.
FT DISULFID 494 505 BY SIMILARITY.
FT DISULFID 499 514 BY SIMILARITY.
FT DISULFID 516 525 BY SIMILARITY.
FT DISULFID 532 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 554 563 BY SIMILARITY.
FT DISULFID 570 580 BY SIMILARITY.
FT DISULFID 575 589 BY SIMILARITY.
FT DISULFID 591 600 BY SIMILARITY.
FT DISULFID 607 618 BY SIMILARITY.
FT DISULFID 612 627 BY SIMILARITY.
FT DISULFID 629 638 BY SIMILARITY.
FT DISULFID 645 655 BY SIMILARITY.
FT DISULFID 650 664 BY SIMILARITY.
FT DISULFID 666 675 BY SIMILARITY.
FT DISULFID 682 693 BY SIMILARITY.
FT DISULFID 687 702 BY SIMILARITY.
FT DISULFID 704 713 BY SIMILARITY.
FT DISULFID 720 730 BY SIMILARITY.
FT DISULFID 725 739 BY SIMILARITY.
FT DISULFID 741 750 BY SIMILARITY.
FT DISULFID 757 768 BY SIMILARITY.
FT DISULFID 762 777 BY SIMILARITY.
FT DISULFID 779 788 BY SIMILARITY.
FT DISULFID 795 806 BY SIMILARITY.
FT DISULFID 800 815 BY SIMILARITY.
FT DISULFID 817 826 BY SIMILARITY.
FT DISULFID 833 844 BY SIMILARITY.
FT DISULFID 838 855 BY SIMILARITY.
FT DISULFID 857 867 BY SIMILARITY.
FT DISULFID 874 885 BY SIMILARITY.
FT DISULFID 879 894 BY SIMILARITY.
FT DISULFID 896 905 BY SIMILARITY.
FT DISULFID 912 923 BY SIMILARITY.
FT DISULFID 917 932 BY SIMILARITY.
FT DISULFID 934 943 BY SIMILARITY.
FT DISULFID 943 949 BY SIMILARITY.
FT DISULFID 988 999 BY SIMILARITY.
FT DISULFID 993 1008 BY SIMILARITY.
FT DISULFID 1010 1019 BY SIMILARITY.
FT DISULFID 1026 1037 BY SIMILARITY.
FT DISULFID 1031 1046 BY SIMILARITY.
FT DISULFID 1048 1057 BY SIMILARITY.
FT DISULFID 1064 1075 BY SIMILARITY.
FT DISULFID 1069 1084 BY SIMILARITY.
```

```
FT DISULFID 1086 1095 BY SIMILARITY.
FT DISULFID 1102 1123 BY SIMILARITY.
FT DISULFID 1117 1132 BY SIMILARITY.
FT DISULFID 1134 1143 BY SIMILARITY.

Query Match 23.2%; Score 292.5; DB 1; Length 2444;
Best Local Similarity 32.3%; Pred. No. 1.3e-16;
Matches 60; Conservative 16; Mismatches 45; Indels 65; Gaps 5;

QY 26 CNPNPCNGCIGLSGLADDSFSCBEP----- 51
DB 833 CAPSPCRNGCECROSEDEYESFCVCPTRGAKGOTCEVDINECVLSPCRHGASCONTHGX 892
QY 52 -----EGFAGPNCSSVVEVASDEEKPESAGPCIPNPNCHNGTCE--ISEAYRGDTFISYV 104
DB 893 RCHCOAGISGRNCE-----TDDCRPNPCNNGSGCTGINTAF----- 931
QY 105 CKCPRGFNGICHOHNIINECEAPPCRNGGICTDLVANYSCECPGEFMRNCQYK---CSGH 161
DB 932 CDLPGEFRGTFCBEDINECASDPCRNGANCCTCDVSYCTCPAGSGIHCENRTPDCTES 991
QY 162 LGIEGG 167
DB 992 SCFNGG 997
```

Search completed: May 23, 2001, 06:19:03
Job time: 367 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:20:50 ; Search time 105.14 Seconds
(without alignments)
246.366 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 1262
Sequence: 1 MKHLVAAMLVLGSLGVPOF.....INMTAENDRWPMVQVTG 221

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_15:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1222	96.8	480	11	035474 mus musculu
2	1161	92.0	480	4	043854 homo sapien
3	402	31.9	463	11	09RIx9
4	381.5	30.2	426	11	09MTs3
5	346	27.4	363	6	077718
6	341	27.0	2352	5	061240
7	332	26.3	406	5	025059
8	317.5	25.2	721	13	091902
9	314.5	24.9	728	13	090656
10	314.5	24.9	2653	5	025253
11	313	24.8	529	5	025058
12	311	24.6	2447	13	013149
13	308.5	24.4	2634	5	09M4T8
14	308.5	24.4	2704	5	09M4T8
15	305.5	24.2	717	13	087357
16	304	24.1	723	4	09ND41
17	303	24.0	192	11	09QWQ1
18	303	24.0	752	13	042374
19	303	24.0	1203	11	006008

20	303	24.0	2470	11	035516	035516 mus musculu
21	301	23.9	585	5	09U0E2	09U0E2 tribolium c
22	299.5	23.7	2531	5	016004	016004 lytechinus
23	299	23.7	723	4	09UJY2	09UJY2 homo sapien
24	298.5	23.7	802	13	057462	057462 brachydanio
25	298	23.6	1404	5	09VB85	09VB85 drosophila
26	297.5	23.6	615	13	057409	057409 brachydanio
27	297	23.5	2471	11	09QW30	09QW30 rattus sp.
28	293.5	23.3	2319	11	09RI72	09RI72 rattus norv
29	292.5	23.2	1530	11	09WU5	09WU5 rattus norv
30	292	23.1	529	5	09V048	09V048 drosophila
31	291	23.1	1218	4	015122	015122 homo sapien
32	291	23.1	1218	4	015816	015816 homo sapien
33	291	23.1	1218	4	014902	014902 homo sapien
34	291	23.1	1227	4	014902	014902 homo sapien
35	290.5	23.0	2281	4	09UPL3	09UPL3 homo sapien
36	290.5	23.0	2321	4	09Y6L8	09Y6L8 homo sapien
37	290.5	23.0	2321	4	09U47	09U47 homo sapien
38	289.5	22.9	861	11	09QW58	09QW58 mus sp. not
39	289	22.9	156	5	026661	026661 strongyloce
40	288.5	22.9	1193	13	090819	090819 gallus gall
41	287.5	22.8	1531	11	088279	088279 rattus norv
42	286	22.7	1218	11	09QXX0	09QXX0 mus musculu
43	286	22.7	1219	11	063722	063722 rattus norv
44	285.5	22.6	589	11	088671	088671 rattus norv
45	285	22.6	1212	13	042347	042347 gallus gall

ALIGNMENTS

RESULT 1
035474 PRELIMINARY; PRT; 480 AA.
ID 035474
AC 035474; 035475;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE INTEGRIN-BINDING PROTEIN DEL1 PRECURSOR.
GN EDL3 OR DEL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC TISSUE-EMBRYO;
RX MEDLINE=96083109; PubMed=9420328;
RA Hidaï C., Zupancic T., Penta K., Mikhail A., Kawana M.,
RA Quertermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
RA Auerbach R., Hogan B.L.M., Snodgrass R., Quertermous T.;
RT "Cloning and characterization of developmental endothelial locus-1: an
RT embryonic endothelial cell protein that binds the alphabeta3 integrin
RT receptor".
RT Genes Dev. 12:21-33(1998).
RL
- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
- INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
- FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
- REGULATION OF VASCULAR MORPHOGENESIS OR REMODELING IN EMBRYONIC
- DEVELOPMENT.
- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND
- SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
- TISSUE SPECIFICITY: EXPRESSED IN ANGIOBLASTS AND EARLY ENDOTHELIAL
- CELLS. BY EMBRYONIC DAY 13.5, ALSO EXPRESSED IN A RETINUAL GROUP
- OF NON-ENDOTHELIAL CELLS INCLUDING CHONDROCYTES AND RETINAL
- NEURONS.
- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. AFTER DAY
- 15.5, EXPRESSION DECREASES AND DISAPPEARS COMPLETELY BY THE TIME
- OF BIRTH.
- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
DR EMBL; AF031524; AAB8585.1; -.

DR EMBL; AF031525; AAB86586.1; -.
DR HSSP; P00740; 1EDM.
DR MGD; MGI:1329025; Ed113.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000421; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001881; -.
DR PFAM; PF00008; EGF_3.
DR PFAM; PF00754; F5_F8_Type_C_2.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR EGF-like domain; Alternative splicing; Signal; Developmental protein;
KW Cell adhesion; Repeat; Vascularization.
FT SIGNAL 1 16
FT CHAIN 17 480
FT DOMAIN 26 59
FT DOMAIN 78 116
FT DOMAIN 123 154
FT DOMAIN 161 311
FT DOMAIN 322 473
FT SITE 96 98
FT DISULFID 26 37
FT DISULFID 31 48
FT DISULFID 50 59
FT DISULFID 78 89
FT DISULFID 83 105
FT DISULFID 107 116
FT DISULFID 158 314
FT DISULFID 301 305
FT DISULFID 319 476
FT VASPLIC 218 221
FT VASPLIC 222 480
SQ SEQUENCE 480 AA; 53740 MM; 4CD91EFE9261714D CRC64;
Query Match 96.8%; Score 1222; DB 11; Length 480;
Best Local Similarity 96.8%; Pred. No. 1.2e-120;
Matches 213; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 1 MKHLVAAMLVLGLSLGVPOFGKGDICNPNCENGICLSGLADSFSCCECEGFAGPCNS 60
DB 1 MKHLVAAMLVLGLSLGVPOFGKGDICNPNCENGICLSGLADSFSCCECEGFAGPCNS 60
OY 61 SVEVASDEEKPVSAGPCIPNPNCHNGGTCEISEAYRGDTFTGYVCKCPRGFGNHCOHNI 120
DB 61 SVEVASDEEKPVSAGPCIPNPNCHNGGTCEISEAYRGDTFTGYVCKCPRGFGNHCOHNI 120
OY 121 NECEAEPCRNNGICTDLVANSCECPGEFPMGRNCQYKCSGHLGEGGIIISNOQTASSNH 180
DB 121 NECEAEPCRNNGICTDLVANSCECPGEFPMGRNCQYKCSGHLGEGGIIISNOQTASSNH 180
OY 181 RALFGLQKWPYPAALNKKGLINAMWTAENDRWPMIOVTV 220
DB 181 RALFGLQKWPYPAALNKKGLINAMWTAENDRWPMIOVTV 220
RESULT 2
O43854 PRELIMINARY; PRT; 480 AA.
AC O43854; O43855;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE * INTEGRIN-BINDING PROTEIN DELL PRECURSOR.
GN DELL
OS Homo sapiens (Human).
OC Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiindae; Homo.

OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND CHARACTERIZATION.
RC TISSUE=EMBRYONIC LUNG;
RA MEDLINE=98083109; PubMed=9420328;
RX Hidel C., Zupancic T.J., Penta K., Mikhail A., Kawana M.,
RA Quertermous E.E., Aoka Y., Fukagawa M., Matsui Y., Platika D.,
RA Auerbach R., Hogan B.L.M., Snodgrass R., Quertermous T.;
RT *Cloning and characterization of developmental endothelial locus-1: an
RT embryonic endothelial cell protein that binds the alphavbeta3 Integrin
RT receptor*;
RL Genes Dev. 12:21-33(1998).
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION WITH THE ALPHA-V BETA3 INTEGRIN RECEPTOR. INHIBITS
CC FORMATION OF VASCULAR-LIKE STRUCTURES. MAY BE INVOLVED IN
CC REGULATION OF VASCULAR MORPHOGENESIS OF REMODELING IN EMBRYONIC
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; LONG ISOFORM (SHOWN HERE) AND
CC SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC EMBL; U70312; AAC02648.1; -.
CC EMBL; U70313; AAC02649.1; -.
DR HSSP; P00740; 11XA.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000421; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR001881; -.
DR PFAM; PF00008; EGF_3.
DR PFAM; PF00754; F5_F8_Type_C_2.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
KW EGF-like domain; Alternative splicing; Signal; Developmental protein;
KW Cell adhesion; Repeat; Vascularization.
FT SIGNAL 1 16
FT CHAIN 17 480
FT DOMAIN 26 59
FT DOMAIN 78 116
FT DOMAIN 123 154
FT DOMAIN 161 311
FT DOMAIN 322 473
FT SITE 96 98
FT DISULFID 26 37
FT DISULFID 31 48
FT DISULFID 50 59
FT DISULFID 78 89
FT DISULFID 83 105
FT DISULFID 107 116
FT DISULFID 158 314
FT DISULFID 301 305
FT DISULFID 319 476
FT VASPLIC 218 221
FT VASPLIC 222 480
SQ SEQUENCE 480 AA; 53765 MM; F7171E23A309FD48 CRC64;
Query Match 92.0%; Score 1161; DB 4; Length 480;
Best Local Similarity 91.4%; Pred. No. 3.3e-114;
Matches 201; Conservative 6; Mismatches 13; Indels 0; Gaps 0;
OY 1 MKHLVAAMLVLGLSLGVPOFGKGDICNPNCENGICLSGLADSFSCCECEGFAGPCNS 60
DB 1 MKRSVAWMLVLGLSLGVPOFGKGDICDPNCPENGICLPLGLADSFSCCECEGFAGPCNS 60
OY 61 SVEVASDEEKPVSAGPCIPNPNCHNGGTCEISEAYRGDTFTGYVCKCPRGFGNHCOHNI 120
DB 61 SVEVASDEEKPVSAGPCIPNPNCHNGGTCEISEAYRGDTFTGYVCKCPRGFGNHCOHNI 120

Y*	121	NECEEPEPCNGICGIDLVANYSCEPGEPMGNCOYKCSGHGIEGIIISNOQITASSNH	180
Db	121	NECEEPEPCNGICGIDLVANYSCEPGEPMGNCOYKCSGHGIEGIIISNOQITASSNH	180
Qy	181	RALEGLQKWPYPYALANKKGLINMTAENDRPMWIOQTV	220
Db	181	RALEGLQKWPYPYALANKKGLINMTAENDRPMWIOQTV	220
RESULT	3		
ID	09RLX9	PRELIMINARY;	PRT; 463 AA.
AC	09RLX9;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)		
DE	MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 LONG FORM.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BAJB/C; TISSUE=MAMMARY GLAND;		
RA	Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.,		
RT	"Lactation-dependent expression of an mRNA splice variant with an exon		
RT	for a multiply O-glycosylated domain of mouse milk fat globule		
RL	glycoprotein MFG-E8."		
RL	Biochem. Biophys. Res. Commun. 254:522-528(1999).		
DR	EMBL: AB021130; BAA35180.1; -		
DR	HSSP: P00740; 11YA.		
DR	INTERPRO: IPR000421; -		
DR	INTERPRO: IPR000561; -		
DR	INTERPRO: IPR001092; -		
DR	INTERPRO: IPR001438; -		
DR	PFAM: PF00008; EGF_2.		
DR	PFAM: PF00754; F5_F8_type_C_2.		
DR	PRINTS: PR00010; EGRBLOOD.		
DR	PROSITE: PS00022; EGF_1; UNKNOWN_2.		
DR	PROSITE: PS01186; EGF_2; 2.		
DR	PROSITE: PS01285; FA58C_1; 2.		
DR	PROSITE: PS01286; FA58C_2; 2.		
SO	SEQUENCE 463 AA; 51269 MW; D119D2BE090E6427 CRC64;		
Query Match			
Best Local Similarity		31.9%; Score 402; DB 11; Length 463;	
Matches		80; Conservative 30; Mismatches 73; Indels 18; Gaps 5	
Qy	23	GDICNPNCENGICISGLADSFSCPEGRAGPNCSSVEVADSEKPTSAQCIPNP	82
Db	25	GFPCSSSLCINGTCITG-QDNDIYICLCEGPTGLVCNE-----TERGPCSPNP	72
Qy	83	CHNGGFCETS-FAYGDDFTIGYVCCQPRGFNGHCOHNTNECEAEPCRNGICITDLVANY	141
Db	73	CYNDACKLVLTLDYQKDIPTETIYICQCPVSGIHCTETLVNYLND---GEYFETTAVPMT	129
Qy	142	SCECPGEF--MGRNCOYKCSGHLGIEGIIISNOQITASSNHRALEGLQKWPYPYALANKK	199
Db	130	AVPTAPAPRPDLISNNLASRCSQGLMGEGALADSQISASSVYMGFMGLQKWPYALANKK	189
Qy	200	GLINMTAENDRPMWIOQTV	220
Db	190	GLINMTAENDRPMWIOQTV	210
RESULT	4		
ID	09WTS3	PRELIMINARY;	PRT; 426 AA.
AC	09WTS3;		
DT	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)		

DE	01-OCT-2000 (TremblRel. 15, Last annotation update)
DF	MILK FAT GLOBULE GLYCOPROTEIN MFG-E8 S.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_Taxid=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BALE/C; TISSUE=MAMMARY GLAND;
RX	MEDLINE=99120894; PubMed=9920772;
RT	Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
RT	"tactation-dependent expression of an mRNA splice variant with an exon
RT	for a multiply O-glycosylated domain of mouse milk fat globule
RT	glycoprotein MFG-E8.";
RL	Biochem. Biophys. Res. Commun. 254:522-528(1999).
DR	EMBL; AB025280; BAA76386.1; -.
DR	HSSP; P00740; IIXA.
DR	INTERPRO; IPR000421; -.
DR	INTERPRO; IPR000561; -.
DR	INTERPRO; IPR001092; -.
DR	INTERPRO; IPR001438; -.
DR	PFAM; PF00008; EGF_2.
DR	PFAM; PF00754; F5_P8_type_C; 2.
DR	PRINTS; PR00010; EGFBLD.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR	PROSITE; PS01186; EGF_2; 2.
DR	PROSITE; PS01285; FA58C_1; 2.
DR	PROSITE; PS01286; FA58C_2; 2.
SO	SEQUENCE 426 AA; 47197 MW; B182FAA20629881A CRC64;
Query Match	30.2%; Score 381.5; DB 11; Length 426;
Best Local Similarity	37.2%; Pred. No. 4e-32;
Matches 74; Conservative 25; Mismatches 49; Indels 51; Gaps	
OY	23 GTCNPNPCENGCIGCLSGADSPSCSECEBGRAGPNCSSVEYASDEKPTNAGCIENP 82
DB	25 GFPCDSSSLCINGCTCTG-QDNDIYCLCEPGEFTGLVCE-----TERGPCSPND 72
OY	83 CHNGGCEITS-LAYGDDPTIGYVCKPRGFCNGCHNHNECEAEPCRRNGIGCTDLVANY 141
DB	73 CYNDAKCLVTLDTQRKDITERYICQCPVYSIGHCTG----- 110
OY	142 SCECEGEFGRMCQYKCSGHLGIEGGIISNCOITVASSNHRALFGLQKWYPYALNKKGL 201
DB	111 -----CSYQLMGEAGALDSQISASVYMGFMGLRMGPDLARLYRTGI 154
OY	202 INAMTAENDRMPWIOYV 220
DB	155 VNAMTASNYDSKPMIOVNL 173
RESULT 5	
ID 077718	PRELIMINARY; PRT; 363 AA.
AC 077718	
DT 01-NOV-1998 (TremblRel. 08, Created)	
DT 01-NOV-1998 (TremblRel. 08, Last sequence update)	
DT 01-MAY-2000 (TremblRel. 13, Last annotation update)	
DE SPMEM-MEMBRANE ASSOCIATED PROTEIN P47 (FRAGMENT).	
OC Equus caballus (Horse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.	
OX NCBI_Taxid=9796;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=TESTIS;	
RA Gentzel M., Toepfer-Petersen E.;	
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AJ010121; CAA09010.1; -.	
DR INTERPRO; IPR000421; -.	
DR INTERPRO; IPR000561; -.	
DR PFAM; PF00008; EGF_1.	

DR PFAM: PF00754; F5_F8-type_C; 2.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01285; FAS5C_1; 1.
KW Sperm.
FT NON_TER 1 1
FT NON_TER 363 363
SQ SEQUENCE 363 AA; 40744 MW; 1F8B6395AF3233BD CRC64;

Query Match 27.4%; Score 346; DB 6; Length 363;
Best Local Similarity 36.2%; Pred. No. 1.8e-28;
Matches 67; Conservative 22; Mismatches 46; Indels 50; Gaps 3;

OY 37 CLSLADDSFSCPEGFGAGPNCSSVEVASDEEKPSTAGIPNCHNGICE-ISMAY 95
DB 2 CLLGDDDLFFYLCPQGFGLICNE-----TEKGPCFPNPNCDGCHVIDSH 50
OY 96 RQDTFIVGVCKPRGFNGIHCOHINNECEAPCRNGICTDLVANYSCPEGFGMRNCQ 155
DB 51 RQDVFTQYICSPRGYTGHCET-----74
OY 156 YKCSGHLGEGGIIISNOQTASSNRALEGLQKMTPTAALNKGILNMTAENDRWP 215
DB 75 --CAMPLEGALADAOISASSVYEGFMGLQRMVPELARLHRTGIVANMTASNDKNPW 132
OY 216 IQTV 220
DB 133 IOVNL 137

RESULT 6
ID 061240 PRELIMINARY; PRT; 2352 AA.
AC 061240:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HRNOTCH PROTEIN.
DE HRNOTCH.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RA Horii S., Saitoh T., Matsumoto M., Makabe K.W., Nishida H.;
RL Dev. Genes Evol. 207:371-380(1997).
DR EMBL: AB001327; BAA2571.1; -.
DR HSSP: P00740; 1EDM.
DR INTERPRO: IPR000152; -.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR000742; -.
DR INTERPRO: IPR000742; -.
DR INTERPRO: IPR000800; -.
DR INTERPRO: IPR001438; -.
DR INTERPRO: IPR001881; -.
DR INTERPRO: IPR002110; -.
DR PFAM: PF00008; EGF_32.
DR PFAM: PF00023; ank; 6.
DR PFAM: PF00066; notch; 3.
DR PRINTS: PR00010; EGFBL00D.
DR PROSITE: PS00010; ASX_HYDROXYL; 18.
DR PROSITE: PS00022; EGF_1; UNKNOWN_28.
DR PROSITE: PS01186; EGF_2; 22.
DR PROSITE: PS01187; EGF_CA; 18.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 2352 AA; 252622 MW; 13DB1C056BBD08D CRC64;

Query Match 27.0%; Score 341; DB 5; Length 2352;
Best Local Similarity 42.1%; Pred. No. 5e-27;
Matches 61; Conservative 17; Mismatches 45; Indels 22; Gaps 4;

OY 26 CNPNPCENGICISGLADDSFSCPEGFGAGPNCSSVEVASDEEKPSTAGIPNCHN 85
DB 720 CATPCQNGGTCTSGI--NSTNCACPAKTYVNE-----TEISPVCNPNCHN 765
OY 86 GGTCEISEAYRGDTFIVGVCKPRGFNGIHCOHINNECEAPCRNGICTDLVANYSC 145
DB 766 GATCQESADY-----LAVYQCEGFRGPTCARDINCVNPKNGGCGTILVPGYQCTC 820
OY 146 PGFMGRNCOY---KCSGHLGIEG 167
DB 821 SQGFTGKDCDTIDDCSSNPLNMG 845

RESULT 7
ID 025059 PRELIMINARY; PRT; 406 AA.
AC 025059:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE FTBROPELLIN III (FRAGMENT).
OS Helicoidaris erythrogramma (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinometridae;
OX Helicoidaris.
OX NCBI_TaxID=7634;
RN [1]
RP SEQUENCE FROM N.A.
RA Bisgrove B.W.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L33862; AAA29996.1; -.
DR HSSP: P00740; 1EDM.
DR INTERPRO: IPR000083; -.
DR INTERPRO: IPR000088; -.
DR INTERPRO: IPR000152; -.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR000742; -.
DR INTERPRO: IPR001010; -.
DR INTERPRO: IPR001438; -.
DR INTERPRO: IPR001881; -.
DR INTERPRO: IPR001947; -.
DR INTERPRO: IPR002049; -.
DR PFAM: PF00008; EGF_7.
DR PFAM: PF01382; Avidin; 1.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00012; ENTYPET.
DR PRINTS: PR00012; ENTYPET.
DR PRINTS: PR00286; CHARYBDTOXIN.
DR PRINTS: PR00287; THIONIN.
DR PRINTS: PR00709; AVIDIN.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_7.
DR PROSITE: PS00022; EGF_1; UNKNOWN_7.
DR PROSITE: PS00572; AVIDIN; 1.
DR PROSITE: PS01186; EGF_2; 6.
DR PROSITE: PS01187; EGF_CA; 5.
KW Glycoprotein; EGF-like domain.
FT NON_TER 1 1
SQ SEQUENCE 406 AA; 43475 MW; EC89CC8ABB7E89A CRC64;

Query Match 26.3%; Score 332; DB 5; Length 406;
Best Local Similarity 34.5%; Pred. No. 6.2e-27;
Matches 60; Conservative 25; Mismatches 51; Indels 38; Gaps 4;

OY 23 GDICNPNCENGICISGLADDSFSCPEGFGAGPNCSSVEVASD-----68
DB 13 GDICNPNCNGAACIDQYND--YECICPPGFTGDCETDIDVCASAPCRNGACVDGVN 70
OY 69 -----BEKPTSAQPCIPNPNCHNGICETSEAYRGDTFIVGVCKPRGFNGIHC 116
DB 71 GYTCNCIPGFDGNCENNNINNECASNPONGVCI-----DGVGFVCTCPGTYGLIC 123
OY 117 QHININECEAPCRNGICTDLVANYSCPEGFGMRNCQ---YKCSGHLGIEG 167

Db 124 ETDIDECASNFCQNGVCTDLVNNYTCDCLAGFTGSCNCEININECASNPCLNG 177

```
RESULT 8
ID 091902 PRELIMINARY; PRT; 721 AA.
AC 091902;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE X-DELTA-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95319507; PubMed=7596411;
RA Henrique D., Adam J., Myat A., Chitnis A., Lewis J., Ish-Horowitz D.;
RT "Expression of a Delta homologue in prospective neurons in the
RT chick.";
RL Nature 375:787-790(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95319503; PubMed=7596407;
RA Chitnis A.B., Henrique D., Lewis J., Ish-Horowitz D., Kinnier C.R.;
RT "Primary neurogenesis in Xenopus embryos regulated by a homologue of
RT the Drosophila neurogenic gene Delta.";
RL Nature 375:761-766(1995).
DR EMBL: LA2229; AAC38017.1; -.
DR HSSP: P00740; 1EDM.
DR INTERPRO: IPR000152; -.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR000742; -.
DR INTERPRO: IPR001438; -.
DR INTERPRO: IPR001774; -.
DR INTERPRO: IPR001881; -.
DR PFAM: PF00008; EGF_8.
DR PFAM: PF01414; DSL; 1.
DR PRINTS: PR00010; EGFBL00D.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 8.
DR PROSITE: PS01187; EGF_CA; 2.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 721 AA; 79922 MW; 9EBDC85C439D194 CRC64;
```

Query Match 25.2%; Score 317.5; DB 13; Length 721;
Best Local Similarity 28.5%; Pred. No. 4e-25;
Matches 70; Conservative 33; Mismatches 66; Indels 77; Gaps 9;

```
QY 26 CNPNCENGICISGLADSFSCBCEPGEFAGPNCSSVEVASDEKPTSAFCI----- 79
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 334 CDANPCKNKGSCSD--LENSYTCSCPPGFYGNKCELSAMTCAD-----GPCFNGGRCA 384
QY 80 -----PNCNGGTCTEISEAYRGDTFTGYCKCPRG 111
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 385 DNPDCGICPCPVYSGFNCCKIDYCSPNCANGARCF-----DLGNSYTCQCEG 437
QY 112 NGIHQNHINECEAPRCNGICITDLVANSCECPGEFMRNCQY--KCSGLGIEGG 168
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 438 SGRNCDDNLDCTSFPCQNGTCTGQDGVNDYSCPCPGYIGNKNSMPTKE -HNPCHNGA 496
QY 169 ISNQITASSNHRAL-----FGLQKWYPTYYAALNKKGLINMTAENDRW 215
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 497 TCHER-----NNRYVCCAGYGANNQCFLLPEKPVVDLTER-----YTEGSGSQFPW 546
QY 216 IOYTVG 221
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 547 IAVCAG 552
```

RESULT 9
ID 090656 PRELIMINARY; PRT; 728 AA.
AC 090656;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE TRANSMEMBRANE PROTEIN C-DELTA-1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPINAL CORD;
RX MEDLINE=95319507; PubMed=7596411;
RA Henrique D., Adam J., Myat A., Chitnis A., Lewis J., Ish-Horowitz D.;
RT "Expression of a Delta homologue in prospective neurons in the
RT chick.";
RL Nature 375:787-790(1995).
DR EMBL: U26590; AAC59689.1; -.
DR HSSP: P00740; 1EDM.
DR INTERPRO: IPR000152; -.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR000742; -.
DR INTERPRO: IPR001438; -.
DR INTERPRO: IPR001774; -.
DR INTERPRO: IPR001881; -.
DR PFAM: PF00008; EGF; 8.
DR PFAM: PF01414; DSL; 1.
DR PRINTS: PR00010; EGFBL00D.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 8.
DR PROSITE: PS01187; EGF_CA; 2.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 728 AA; 79661 MW; 93B2D666D2388B7 CRC64;

Query Match 24.9%; Score 314.5; DB 13; Length 728;
Best Local Similarity 29.1%; Pred. No. 8.3e-25;
Matches 72; Conservative 24; Mismatches 74; Indels 77; Gaps 7;

```
QY 26 CNPNCENGICISGLADSFSCBCEPGEFAGPNCSSVEVASDEKPTSAFCI----- 85
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 339 CDANPCKNKGSCSD--LENSYTCSCPPGFYGNKCE-----LSAMTCADGPCFN 384
QY 86 GG-----TCTEISEAYRG-----DTFTGYCKCPRGNG 113
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 385 GGRCDNPDGCGISCPGLYSGFNCCKIDYCSPNCANGAQCVDLGNSTICQCAAGTG 444
QY 114 IHQNHINECEAPRCNGICITDLVANSCECPGEFMRNCQYKSG-----HLGIEGG 167
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 445 RHCDNDVDDCAFPVNGTCTGQDGVNDYSCPCPGYIGNKNSMPTKE--RCHNPGCHNGATCH 504
QY 168 ISNQITASSNHRALFGLQKWYPTYYAALNKKGL-----NMTAENDRW 214
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 505 ERSNRYVECCAG-----YGLNCOFLLPEPPGVIVDTFTEKTEGONSQFP 552
QY 215 WIOYTVG 221
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 553 WIAVCAG 559
```

```
RESULT 10
ID 025253 PRELIMINARY; PRT; 2653 AA.
AC 025253;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
```

01-OCT-2000 (Tremblrel, 15, Last annotation update)
DE NOTCH HOMOLOG SCALLOPED WINGS (SCL).
GN SCL.
OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Coleoptera; Calliphoridae; Lucilia.
OX NCBI_TaxID=7375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SS SEEKING;
RA MEDLINE=96400928; PubMed=8807304;
RX Davies A.G., Game A.Y., Chen Z., Williams T.J., Goodall S., Yen J.L.,
RA McKenzie J.A., Batterham P.;
RT "Scalloped wings is the Lucilia cuprina Notch homologue and a
RT candidate for the modifier of fitness and asymmetry of diazot
RL resistance." Genetics 143:1321-1337(1996).
RL [2]
RN SEQUENCE OF 39-265 FROM N.A.
RC STRAIN-SS SEEKING;
RA Chen Z., McKenzie J.A., Batterham P.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 39-265 FROM N.A.
RC STRAIN-SS SEEKING;
RA Chen Z., McKenzie J.A., Batterham P.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58977; AAC36151.1; -;
DR EMBL; AF032672; AAC36152.1; -;
DR EMBL; AF032670; AAC36152.1; JOINED.
DR EMBL; AF032671; AAC36152.1; JOINED.
DR EMBL; AF032673; AAC36153.1; -;
DR HSSP; P00740; 1EDM.
DR INTERPRO; IPR000152; -;
DR INTERPRO; IPR000561; -;
DR INTERPRO; IPR000742; -;
DR INTERPRO; IPR000800; -;
DR INTERPRO; IPR001438; -;
DR INTERPRO; IPR001881; -;
DR INTERPRO; IPR002110; -;
DR PFAM; PF00008; EGF; 36.
DR PFAM; PF00023; ank; 6.
DR PFAM; PF00066; notch; 3.
DR PRINTS; PS00010; EGFBLDOD.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; UNKNOWN_34.
DR PROSITE; PS01186; EGF_CA; 21.
DR PROSITE; PS01187; EGF_CA; 28.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 2653 AA; 285928 MW; 6AF2A058FE6C329 CRC64;

Query Match 24.9%; Score 314.5; DB 5; Length 2653;
Best Local Similarity 32.6%; Pred. No. 3.6e-24;
Matches 60; Conservative 24; Mismatches 63; Indels 37; Gaps 6;

QY 26 CNPNCEGNGICISGLADSFSCCEPGEFAGPNCSSVVEVASDEKFRSPACPCIPNCHN 85
DB 203 CDDNCKRGHGTCLN--THGSYQCMCPAGYTKNCES-----KYVPCSPSCON 248
QY 86 GATCEISAYAGDFIFGVYVCKPRGFNGIHCOHNIINECEAEPCRNNGICTDLVANYCEC 145
DB 249 GGTCT-----STGLYTECKPCEGYCKNCEQNIIDCPGHLCONGTCTDGINSTYHCAC 301
QY 146 PGEFNGRCQYKCSGHLIEGGIISNDQITASSNHRALFGLQKWPYYAALNKKGLINAW 205
DB 302 PPNYTGENCE--KDVDECAIRBSVCQNGATCTNSGGS-----YSCI-----CVNGW 345
QY 206 TAAE 209
DB 346 TGPD 349

RESULT 11
Q25058
ID 025058 PRELIMINARY; PRT; 529 AA.
AC Q25058;
DT 01-NOV-1996 (Tremblrel, 01, Created)
DT 01-NOV-1996 (Tremblrel, 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel, 15, Last annotation update)
DE FIROPELLIN IA (FRAGMENT).
OS Helicoidaris erythrogramma (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Echinoidae; Echinacea; Echinoida; Echinometridae;
OX NCBI_TaxID=7634;
RN [1]
RP SEQUENCE FROM N.A.
RA Bisgrove B.W.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L33861; AAA29995.1; -;
DR HSSP; P08709; 1BF9.
DR INTERPRO; IPR000083; -;
DR INTERPRO; IPR000088; -;
DR INTERPRO; IPR000152; -;
DR INTERPRO; IPR000561; -;
DR INTERPRO; IPR000742; -;
DR INTERPRO; IPR001010; -;
DR INTERPRO; IPR001438; -;
DR INTERPRO; IPR001881; -;
DR INTERPRO; IPR001947; -;
DR INTERPRO; IPR002049; -;
DR PFAM; PF00008; EGF; 10.
DR PFAM; PF01382; Avidin; 1.
DR PRINTS; PRO0010; EGFBLDOD.
DR PRINTS; PRO0011; EGFAMININ.
DR PRINTS; PRO0012; ENTPEI.
DR PRINTS; PRO0286; CHARYBDTOXIN.
DR PRINTS; PRO0287; THIONIN.
DR PRINTS; PRO0709; AVIDIN.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_8.
DR PROSITE; PS00022; EGF_1; UNKNOWN_11.
DR PROSITE; PS00577; AVIDIN; 1.
DR PROSITE; PS01186; EGF_2; 10.
DR PROSITE; PS01187; EGF_CA; 7.
KW Glycoprotein; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 529 AA; 55543 MW; D4AD958FCF9ACB5A CRC64;

Query Match 24.8%; Score 313; DB 5; Length 529;
Best Local Similarity 30.2%; Pred. No. 8.3e-25;
Matches 65; Conservative 24; Mismatches 50; Indels 76; Gaps 6;

QY 24 DICPNPCENGICISGLADSFSCCEPGEFAGPNCSS-----VVEVASD 68
DB 175 DECAAPCEGNGICIGAV--NGYTCNACASGYGTGNCETIEDCASPCLNGQCIEHWNG 232
QY 69 EEPRTSAG-----PCIPNCHNGTCEISAYAGDFIFGVYVCKPRGFNGIHCO 117
DB 233 YTCQCAAGFTGVLCEFDIDECASDPQONGVCT-----DIVNGTICGVGFTGSDCE 285
QY 118 HNIINECEAEPCRNNGICTDLV-----A 139
DB 286 TINIINECASGCGNGGCVGVNGVFCQCPNRYGTGYCEISLDACSMPCQNGATCVNNGA 345
QY 140 NYSCECPGEFMRNCOY---KCSGHLGIEGGIISN 171
DB 346 NYICECPGEFAGQCEIIDINECASLPQNGGVGIN 380
RESULT 12
QY 013149 PRELIMINARY; PRT; 2447 AA.
ID 013149
AC 013149;

```

DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE NOTCH 2 (FRAGMENT).
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Retrodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T., Trowsdale J.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB004829; BAA20535.1; -.
DR HSSP; P00740; 1EDM.
DR INTERPRO; IPR000083; -.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR000800; -.
DR INTERPRO; IPR001010; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001881; -.
DR INTERPRO; IPR001947; -.
DR INTERPRO; IPR002049; -.
DR INTERPRO; IPR002110; -.
DR PFAM; PF00008; EGF; 35.
DR PFAM; PF00023; ank; 6.
DR PFAM; PF00066; notch; 3.
DR PRINTS; PRO0010; EGFBL00D.
DR PRINTS; PRO0011; EGFBLMIN.
DR PRINTS; PRO0012; ENTPEI.
DR PRINTS; PRO0286; CHARYBDTOXIN.
DR PRINTS; PRO0287; THIONIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; UNKNOWN_35.
DR PROSITE; PS01186; EGF_2; 29.
DR PROSITE; PS01187; EGF_CA; 20.
KM Glycoprotein; EGF-like domain.
FT NON_TER 1
SQ SEQUENCE 2447 AA; 262542 MW; 1A8E2A372A085D84 CRC64;

Query Match
Best Local Similarity 24.6%; Score 311; DB 13; Length 2447;
Matches 66; Conservative 19; Mismatches 42; Indels 76; Gaps 8;

QY 24 DICNPNCENGICLSGLAD--SFSCCEBGFAGPNCSSVVEASDEEKPPTSAGPCIPN 81
DB 867 DDCSPNCLNGSCV-----DDVGSFSCCEBGFAGPNCSSVVEASDEEKPPTSAGPCIPN 910
QY 82 PCNNGTCEISEARGTFTFYGVCKPCRFNGHICHOINT-----CASQ 910
DB 911 PCRNAGTCA-----DYVNSFVCECRLEGEDILCDHNLICTESSCLNNGTCTIDINTF 963
QY 121 -----NECEAEPCRNNGTCTDLVANYSCCEBGFAGPNCSSVVEASDEEKPPTSAGPCIPN 155
DB 964 SCRLPGEFTGFCEYEGNECHSCQCKNGGCTDGLGTGRCTCPAGYNGQNCQNVNLCRQ 1023
QY 156 YKCSGHLGIEGIIISNOQITASSNH 180
DB 1024 VRCH-----NGSGCSHTGATSWTCH 1043

RESULT 13
Q9W4T8 PRELIMINARY; PRT; 2634 AA.
AC Q9W4T8;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE N GENE PRODUCT.

```

```

GN N.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sulten G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champs M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasno P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshell A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter R., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT *The genome sequence of Drosophila melanogaster.*;
RL Science 287:2185-2195 (2000).
DR EMBL; AF003426; AAF45848.1; -.
DR HSSP; P00740; 1IYA.
DR FLYBASE; FBgn004647; N.
DR INTERPRO; IPR000152; -.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR000742; -.
DR INTERPRO; IPR000800; -.
DR INTERPRO; IPR001438; -.
DR INTERPRO; IPR001881; -.
DR INTERPRO; IPR002049; -.
DR INTERPRO; IPR002110; -.
DR PFAM; PF00008; EGF; 36.
DR PFAM; PF00023; ank; 6.
DR PFAM; PF00066; notch; 3.
DR PRINTS; PRO0010; EGFBL00D.
DR PRINTS; PRO0011; EGFBLMIN.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 28.
DR PROSITE; PS01187; EGF_CA; 21.
SQ SEQUENCE 2634 AA; 281558 MW; C9D52C20150BBD3D CRC64;

Query Match
Best Local Similarity 24.4%; Score 308.5; DB 5; Length 2634;
Matches 66; Conservative 19; Mismatches 42; Indels 76; Gaps 8;

```

```

Matches 53; Conservative 21; Mismatches 51; Indels 33; Gaps 2;
QY 24 DICNPNPCENGICLSGLADDSFSCBCEPBGFPAGPNCSSVVEVA----- 66
Db 179 NLCASSPCRNAGCTALAGSSSFTCSCPGFTGDTCSYDIECQSNPKYGGTCVNTGHS 238
QY 67 -----SDEKPTSAQCIPNPNCHNGTCEISEAVRGDFIIGYVCKPGEFGJHCQ 117
Db 239 YQCMCTGYTGKDCDPTKYPKCPSPCQNGICR-----SNGLSYECKCPKGEFGKNC 291
QY 118 HNIKECAEPORNGICITDLVANYSCBCEPGEFGNRNQ 155
Db 292 QNYDDCLGHLCQNGCTCIDGIDSYTCRCPPNFTGRFCQ 329

RESULT 14
097458 PRELIMINARY; PRT; 2704 AA.
AC 097458;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE EG:163A10.2 PROTEIN.
GN EG:163A10.2 OR EG:140G11.1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Borkova D., Minana B., Kafatos F.C.;
RT "Sequencing the distal x chromosome of Drosophila melanogaster.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RE [2]
RA Benos P.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035395; CAB37611.1;
DR EMBL; AL035436; CAB37610.1; JOINED.
DR EMBL; AL035395; CAB37610.1; JOINED.
DR HSSP; P00740; 1EDM.
DR INTERPRO; IPR000152;
DR INTERPRO; IPR000561;
DR INTERPRO; IPR000742;
DR INTERPRO; IPR000800;
DR INTERPRO; IPR001438;
DR INTERPRO; IPR001881;
DR INTERPRO; IPR002049;
DR INTERPRO; IPR002110;
DR PFAM; PF00008; EGF_36.
DR PFAM; PF00023; ank_6.
DR PFAM; PF00066; notch_3.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGFAMININ.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; UNKNOWN_34.
DR PROSITE; PS01186; EGF_2; 28.
DR PROSITE; PS01187; EGF_CA; 21.
KM Glycoprotein; EGF-like domain.
SQ SEQUENCE 2704 AA; 288894 MW; 554298A25C9EBAE CRC64;

```

```

Db 239 YQCMCTGYTGKDCDPTKYPKCPSPCQNGICR-----SNGLSYECKCPKGEFGKNC 291
QY 118 HNIKECAEPORNGICITDLVANYSCBCEPGEFGNRNQ 155
Db 292 QNYDDCLGHLCQNGCTCIDGIDSYTCRCPPNFTGRFCQ 329

RESULT 15
P87357 PRELIMINARY; PRT; 717 AA.
ID P87357
AC P87357;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE DELTAD TRANSMEMBRANE PROTEIN PRECURSOR.
GN DLD OR DELTAD.
OS Brachydanio rerio (Zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97346722; PubMed-9203139;
RA Dornseiffer P., Takke C., Campos-Ortega J.A.;
RT "Overexpression of a zebrafish homologue of the Drosophila neurogenic
RT gene Delta perturbs differentiation of primary neurons and somite
RT development.";
RL Mech. Dev. 63:159-171(1997).
DR HSSP; P00740; 1EDM.
DR EMBL; Y11760; CAA72425.1;
DR ZFIN; ZDB-GENE-990415-47; dld.
DR INTERPRO; IPR000152;
DR INTERPRO; IPR000561;
DR INTERPRO; IPR000742;
DR INTERPRO; IPR001438;
DR INTERPRO; IPR001774;
DR INTERPRO; IPR001881;
DR PFAM; PF00008; EGF_8.
DR PFAM; PF01414; DSL_1.
DR PRINTS; PR00010; EGFBL00D.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_8.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 2.
KM Signal; Transmembrane; Glycoprotein; EGF-like domain.
FT SIGNAL 1
FT CHAIN 1
FT POTENTIAL.
FT DELTAD TRANSMEMBRANE PROTEIN.
SQ SEQUENCE 717 AA; 79061 MW; 9C5A0162504593E4 CRC64;

```

```

Query Match 24.2%; Score 305.5; DB 13; Length 717;
Best Local Similarity 27.6%; Pred. No. 7.3e-24;
Matches 70; Conservative 25; Mismatches 70; Indels 89; Gaps 8;
QY 26 CNPNCEGNGICLSGLADDSFSCBCEPBGFPAGPNCSSVVEVASEDEKPTSAQCIPNPNCHN 85
Db 330 CGGSPCRNAGCTCTALAGSSSFTCSCPGFTGDTCSYDIECQSNPKYGGTCVNTGHS 375
QY 86 GG-----TCBISBAYRG-----DTFYGVCCKPGEFG 113
Db 376 GGHCADNPEGGYFCQCPMGYAGFENCEKKIDHCSSNPNCAOCDDLDVDSYLCQCPBEGFTG 435
QY 114 IHQCNINIECAEPORNGICITDLVANYSCBCEPGEFGNRNC----- 154
Db 436 THCEDNIDECATYPCQNGCTCIDGIDSYTCRCPPYTGKNTSAVNKCLNPNCHNGATC 495
QY 155 -----QYKSGHUGTEGGIISNQITASSNHRA---LFGLOKMYRYAALAKKGLINAWTA 207
Db 496 EMDNRIVYACLPYTG---RNCQFLLENPFGQAIIVEGADRRSY-----E 538
QY 208 AENDRPWIVQTVG 221

```


Wed May 23 07:31:14 2001

us-09-237-981-29.rspt

Page 9

Db 539 EDDGGFPWTAVCAG 552

Search completed: May 23, 2001, 06:20:52
Job time: 446 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:15:17 ; Search time 75.05 Seconds

(without alignments)
168,329 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 1262
Sequence: 1 MKHLVAAWLVGLSLGVPOF.....INMTAAENDRWPMIOVTG 221

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.0A01.*
1: /SID56/gcgdata/geneseq/AI1980.DAT.*
2: /SID56/gcgdata/geneseq/AI1981.DAT.*
3: /SID56/gcgdata/geneseq/AI1982.DAT.*
4: /SID56/gcgdata/geneseq/AI1983.DAT.*
5: /SID56/gcgdata/geneseq/AI1984.DAT.*
6: /SID56/gcgdata/geneseq/AI1985.DAT.*
7: /SID56/gcgdata/geneseq/AI1986.DAT.*
8: /SID56/gcgdata/geneseq/AI1987.DAT.*
9: /SID56/gcgdata/geneseq/AI1988.DAT.*
10: /SID56/gcgdata/geneseq/AI1990.DAT.*
11: /SID56/gcgdata/geneseq/AI1991.DAT.*
12: /SID56/gcgdata/geneseq/AI1992.DAT.*
13: /SID56/gcgdata/geneseq/AI1993.DAT.*
14: /SID56/gcgdata/geneseq/AI1994.DAT.*
15: /SID56/gcgdata/geneseq/AI1995.DAT.*
16: /SID56/gcgdata/geneseq/AI1996.DAT.*
17: /SID56/gcgdata/geneseq/AI1997.DAT.*
18: /SID56/gcgdata/geneseq/AI1998.DAT.*
19: /SID56/gcgdata/geneseq/AI1999.DAT.*
20: /SID56/gcgdata/geneseq/AI2000.DAT.*
21: /SID56/gcgdata/geneseq/AI2001.DAT.*
22: /SID56/gcgdata/geneseq/AI2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1262	100.0	221	20	W94684
2	1257	99.6	221	18	W10366
3	1242	98.4	480	18	W10364
4	1242	98.4	480	20	W94683
5	1133.5	89.8	513	20	W94685
6	1132.5	89.7	481	18	W10365
7	542	42.9	102	18	W10370
8	369.5	29.3	426	21	Y94454
9	318	25.2	2471	20	Y06816
10	314.5	24.9	727	18	W11719
11	314.5	24.9	728	21	Y79029

12	314.5	24.9	740	18	W00876
13	311.5	24.7	721	21	Y79030
14	304	24.1	702	18	W18349
15	304	24.1	702	20	W75495
16	304	24.1	723	18	W18353
17	304	24.1	723	20	W94498
18	304	24.1	723	20	W75492
19	304	24.1	723	21	B00172
20	304	24.1	723	21	B34322
21	304	24.1	723	21	B34328
22	304	24.1	723	21	Y83227
23	304	24.1	723	22	B53064
24	298.5	23.7	321	20	W94697
25	298	23.6	1404	21	Y59600
26	297	23.5	512	20	W75494
27	297	23.5	520	18	W18348
28	297	23.5	723	21	Y79032
29	293.5	23.3	1530	21	B28151
30	291.5	23.1	722	18	W11720
31	291	23.1	1010	20	W87896
32	291	23.1	1036	18	W18351
33	291	23.1	1187	18	W18352
34	291	23.1	1208	19	W40827
35	291	23.1	1218	17	W05833
36	291	23.1	1218	19	W44301
37	291	23.1	1218	21	W87894
38	291	23.1	1218	21	Y59597
39	290.5	23.0	2321	19	W49698
40	290	23.0	1218	18	W18354
41	289	22.9	1404	14	R38304
42	288.5	22.9	694	21	B42900
43	288.5	22.9	1193	17	W05835
44	288.5	22.9	1193	21	Y59599
45	288	22.8	157	18	W11730

ALIGNMENTS

RESULT 1	W94684	W94684 standard; Protein; 221 AA.
AC	W94684.	
XX		
DF	05-MAY-1999	(first entry)
XX		
DE	Truncated murine Del-1 protein.	
XX		
KW	Del-1; developmentally-regulated endothelial cell locus 1; cancer;	
KW	discoidin I; factor VIII-like domain; epidermal growth factor; EGF;	
KW	diabetic retinopathy; rheumatoid arthritis; endometriosis;	
KW	angiogenesis.	
XX		
OS	Mus sp.	
XX		
PN	US5877281-A.	
XX		
PD	02-MAR-1999.	
XX		
PF	05-JUN-1996;	96US-0659235.
XX		
PR	05-JUN-1996;	96US-0659235.
XX		
PR	07-JUN-1995;	95US-0480229.
XX		
PA	(PROG-) PROGENITOR INC.	
XX	(UTVA-) UNIV VANDERBILT.	
PI	Hogan B, Quarterman T, Snodgrass HR, Zupancic TJ;	
DR	WPI: 1999-189720/16.	
DR	N-PSDB: X18507.	
XX		

PT Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Claim 3; Column 71-72; 73pp; English.
 XX
 CC The present sequence is truncated murine developmentally-regulated
 CC endothelial cell locus 1 (Del-1). The protein has epidermal growth factor
 CC (EGF) like domains and discoidin I/factor VIII-like domains. The Del-1
 CC proteins have an inhibitory effect on angiogenesis (blood vessel growth),
 CC this activity may be useful clinically to prevent neovascularisation of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischaemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC Integrin, and is an apoptosis inducer.
 CC
 XX Sequence 221 AA;
 SQ

Query Match 100.0%; Score 1262; DB 20; Length 221;
 Best Local Similarity 100.0%; Pred. No. 1.9e-75;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKHLVAAMLVLGSLGVPOFGKGDICNPNCENGICLSGLADDSFSCPEGFPAGNCS 60
 Db 1 mkhlvaawllvgslgvpqfgkgticnpncengiclsgladdsfscpegfagpncs 60
 QY 61 SVEVVASDEEKPTSGAPCIPNPNCHNGGTCEISEAYRGDTFTGYCKPCRFNGICQHNT 120
 Db 61 svvevasdeekptsgapcnpncnggtceiseayrgdtftgyckpcrpfngicqhn1 120
 QY 121 NECEAEPCRNNGICITDLYANVSCCEPGEFMRNCOYKCSGHLGIGGIIISNOQTASSNH 180
 Db 121 necaeepcrngicitdlyanyscepgelfmrncqkcsghlgigglisnqitaasnh 180
 QY 181 RALFGLQKWPYPAALNKKGLINAWTAENDRWPIQVTG 221
 Db 181 ralfglqkwyppyaalnkkglinawtaendrwpvigtvg 221

RESULT 2
 W10366
 ID W10366 standard; Protein: 221 AA.
 XX
 AC W10366;
 XX
 DT 03-MAY-1997 (first entry)
 XX
 DE Murine Del-1 truncated minor protein.
 XX
 KW Del-1: developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KW cardiac ischaemia; stroke; vascular disease; wound healing;
 KW vulnerable; bone formation; diagnosis; therapy.
 XX
 OS Mus sp.
 XX
 FH Key
 FT 1..23 Location/Qualifiers
 FT Peptide /label= Sig_Peptide
 FT Protein /label= Sig_Peptide
 FT Domain /label= Mat_Protein
 FT 26..59 /label= EGF-1
 FT /note= "epidermal growth factor-like domain 1"
 FT 78..116 /label= EGF-2
 FT /note= "epidermal growth factor-like domain 2"
 FT Domain 123..154

FT /label= EGF-3
 FT /note= "epidermal growth factor-like domain 3"
 FT 158..221 /label= Discoidin-1
 FT /label= Discoidin-1
 FT /note= "truncated discoidin I/factor VIII-like
 domain 1"
 FT
 PN W09640769-A1.
 PD 19-DEC-1996.
 XX
 PF 05-JUN-1996; 96WO-US09456.
 XX
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNTV VANDERBILT.
 XX
 PI Hogan B. Quaternous T, Snodgrass HR, Zupancic TJ;
 XX
 DR WPI: 1997-052233/05;
 DR N-PSDB; T47339.
 XX
 PT New developmentally regulated endothelial cell locus-1 (del-1) gene
 PT - used to develop prods. for the diagnosis and treatment of cancer
 PT and conditions involving abnormal angiogenesis
 XX
 PS Example: Fig 12; 137pp; English.
 CC A truncated version (W10366) of murine developmentally-regulated
 CC endothelial cell locus-1 (del-1) gene product (see also W10364)
 CC contains a signal peptide, all 3 EGF-like domains but only a
 CC partial N-terminal discoidin I/factor VIII-like domain (about 40%).
 CC It is the product of a murine del-1 minor sequence (T47339). This
 CC transcript was cloned only from mouse embryonic libraries, but was
 CC verified through cloning of several independent cDNAs.
 CC
 XX Sequence 221 AA;
 SQ

Query Match 99.6%; Score 1257; DB 18; Length 221;
 Best Local Similarity 99.5%; Pred. No. 4e-75;
 Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKHLVAAMLVLGSLGVPOFGKGDICNPNCENGICLSGLADDSFSCPEGFPAGNCS 60
 Db 1 mkhlvaawllvgslgvpqfgkgticnpncengiclsgladdsfscpegfagpncs 60
 QY 61 SVEVVASDEEKPTSGAPCIPNPNCHNGGTCEISEAYRGDTFTGYCKPCRFNGICQHNT 120
 Db 61 svvevasdeekptsgapcnpncnggtceiseayrgdtftgyckpcrpfngicqhn1 120
 QY 121 NECEAEPCRNNGICITDLYANVSCCEPGEFMRNCOYKCSGHLGIGGIIISNOQTASSNH 180
 Db 121 necaeepcrngicitdlyanyscepgelfmrncqkcsghlgigglisnqitaasnh 180
 QY 181 RALFGLQKWPYPAALNKKGLINAWTAENDRWPIQVTG 221
 Db 181 ralfglqkwyppyaalnkkglinawtaendrwpvigtvg 221

RESULT 3
 W10364
 ID W10364 standard; Protein: 480 AA.
 XX
 AC W10364;
 XX
 DT 03-MAY-1997 (first entry)
 XX
 DE Mouse developmentally-regulated endothelial cell locus-1 protein.
 XX
 KW Del-1: developmentally-regulated endothelial cell locus-1;
 KW signal transduction; cancer; tumour marker; angiogenesis;
 KW

```

KW diabetic retinopathy; rheumatoid arthritis; endometriosis;
KM cardiac ischemia; stroke; vascular disease; wound healing;
KM vulnerability; bone formation; diagnosis; therapy.
XX
OS Mus sp.
FH Location/Qualifiers
FT Key 1..23
FT Peptide /label= Sig-peptide
FT Protein 24..480
FT /label= Mat_protein
FT Domain 26..59
FT /label= EGF-1
FT /note= "epidermal growth factor-like domain 1"
FT Domain 79..116
FT /label= EGF-2
FT /note= "epidermal growth factor-like domain 2"
FT Domain 123..154
FT /label= EGF-3
FT /note= "epidermal growth factor-like domain 3"
FT Domain 158..314
FT /label= Discoidin-1
FT /note= "discoidin I/factor VIII-like domain 1"
FT Domain 319..476
FT /label= Discoidin-2
FT /note= "discoidin I/factor VIII-like domain 2"
XX
PN W09640769-A1.
XX
PD 19-DEC-1996.
XX
PE 05-JUN-1996; 96MO-US09456.
XX
PR 07-JUN-1995; 95US-0480229.
XX
PA (PROG-) PROGENITOR INC.
XX (UYVA-) UNIV VANDERBILT.
XX
PI Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
XX
DR WPI; 1997-052233/05.
XX N-PSDB; T47338.
XX
PT New developmentally regulated endothelial cell locus-1 (del-1) gene
PT - used to develop prods. for the diagnosis and treatment of cancer
PT and conditions involving abnormal angiogenesis
XX
PS Claim 3; Fig 6; 137pp; English.
XX
CC Murine Del-1 (W10364) is the polypeptide product of the murine
CC developmentally-regulated endothelial cell locus-1 (del-1) gene
CC (T47338). It shows 94% amino acid homology to the human Del-1
CC protein (W10365). Structurally, members of this novel gene family
CC contain 3 EGF-like domains and 2 discoidin I/factor VIII-like
CC domains. Del-1 is expressed in endothelial and certain tumour
CC cells. Its ability to inhibit vascular formation allows its used
CC as an anti-angiogenic agent. It can be used as a tumour marker,
CC to identify Del-1 binding partners, and to modulate endothelial
CC cell growth and blood vessel formation. Recombinant Del-1 can be
CC produced in transformed host cells utilising vectors incorporating
CC del-1 nucleic acids.
XX
SQ Sequence 480 AA;

```

Query Match 98.4%; Score 1242; DB 18; Length 480;
 Best Local Similarity 98.2%; Pred. No. 7.4e-74;
 Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

1 MKHLVAAMLIVGLGVDFGKGDICNPNCENGIGICLGLADSFSCCEPGEFAGPNC 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 mkhlvaawmlivgslgyvdfgkdgicnppncengigiclsgladdsfsccepfagpnc 60

```

```

QY 61 SYVEVASDEKFTSGPCPLPNPCHNGTCEISEAVRGDTFIGYCKCRPGNGIHCHQNI 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 SYVEVASDEEKPTSGPCPLPNPCHNGTCEISEAVRGDTFIGYCKCRPGNGIHCHQNI 120
QY 121 NECEAEPCRNNGICTDLYANVSCCEPGFMGRNCOYKSGHGLEGGIISMOQTASSNH 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 neceaeprngicdclvaansceepgefmgncqykcsghlglegglisnqqltaasn 180
QY 181 RALFGLQKWYPYAAALNKKGLINAWTAAEENDRWPMTOVTV 220
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 ralfglqkwpyyarlknkglinawtaaeendrwpmwidiql 220

```

RESULT 4
 W94683
 ID W94683 standard; Protein; 480 AA.
 AC W94683;
 DT 05-MAY-1999 (first entry)
 DE Murine Del-1 protein.
 XX
 XX Del-1; developmentally-regulated endothelial cell locus 1; cancer;
 KM discoidin I; factor VIII-like domain; epidermal growth factor; EGF;
 KM diabetic retinopathy; rheumatoid arthritis; endometriosis;
 KM angiogenesis.
 XX
 OS Mus sp.
 XX
 PN US5877281-A.
 XX
 PD 02-MAR-1999.
 XX
 PE 05-JUN-1996; 96US-0659235.
 XX
 PR 05-JUN-1996; 96US-0659235.
 PR 07-JUN-1995; 95US-0480229.
 XX
 PA (PROG-) PROGENITOR INC.
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Hogan B, Queternous T, Snodgrass HR, Zupancic TJ;
 XX
 DR WPI; 1999-189720/16.
 DR N-PSDB; X18506.
 XX
 PT Del-1 polypeptide sequences - useful for treatment of cancer,
 PT diabetic retinopathy, rheumatoid arthritis and endometriosis
 XX
 PS Claim 1; Fig 3; 73pp; English.
 XX
 CC The present sequence is murine developmentally-regulated endothelial
 CC cell locus 1 (Del-1). The protein has epidermal growth factor (EGF) like
 CC domains and discoidin I/factor VIII-like domains. The Del-1 proteins
 CC have an inhibitory effect on angiogenesis (blood vessel growth), this
 CC activity may be useful clinically to prevent neovascularisation of
 CC tissues such as tumour nodules and prevention of metastases. The anti-
 CC angiogenic activity of Del-1 may be used to treat abnormal conditions
 CC that result from angiogenesis, including cancer, diabetic retinopathy,
 CC rheumatoid arthritis and endometriosis. Since Del-1 promotes
 CC angiogenesis it can be used to treat cardiac ischemia, thrombotic
 CC stroke, wound healing and peripheral vascular disease. Del-1 is also
 CC useful for promoting bone formation. Del-1 binds to alpha V beta 3
 CC integrin, and is an apoptosis inducer.
 XX
 SQ Sequence 480 AA;

Query Match 98.4%; Score 1242; DB 20; Length 480;
 Best Local Similarity 98.2%; Pred. No. 7.4e-74;
 Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DR	WPI: 11997-052233/05.
DR	N-PSDB: T47343.
PT	New developmentally regulated endothelial cell locus-1 (del-1) gene
PT	- used to develop prods. for the diagnosis and treatment of cancer
PT	and conditions involving abnormal angiogenesis
XX	
PS	Claim 4; Fig 6; 137pp; English.
XX	
CC	Human Del-1 (W10365) is the polypeptide product of the human
CC	developmentally-regulated endothelial cell locus-1 (del-1) gene
CC	(T47343). It shows 94% amino acid homology to the mouse Del-1
CC	protein (W10364). Structurally, members of this novel gene family
CC	contain 3 EGF-like domains and 2 discoidin I/factor VII-like
CC	domains. Del-1 is expressed in endothelial and certain tumour
CC	cells. Its ability to inhibit vascular formation allows its used
CC	as an anti-angiogenic agent. It can be used as a tumour marker,
CC	to identify Del-1 binding partners, and to modulate endothelial
CC	cell growth and blood vessel formation. Recombinant Del-1 can be
CC	produced in transformed host cells utilizing vectors incorporating
CC	del-1 nucleic acids.
XX	
XX	
SQ	Sequence 481 AA;
Query Match	89.7%; Score 1132.5; DB 18; Length 481;
Best Local Similarity	90.0%; Pred. No. 9.9e-67;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;	
QY	1 MKHLVAAMLVLTGLSGVPOFGKGDICNPNPCNGIGICSLGADDSFSCPCPFAGPNCS 60
DB	1 MKRISAVVLLVGLISGVPTQIGYDIDCDPDCPENGGLICLPJLVAGSFSCCPGDFCDPCS 60
QY	61 SVEEVAASDEKETSAGCPIPNPCNHNGTCEISEARGDTFTIGYVCKPCPGFNGIHCQNI 120
DB	61 SVEEVAASDEEPTSGPCTPNCPCNHNGTCEISEARGDTFTIGYVCKPCPGFNGIHCQNI 120
QY	121 NECEAEPCRNIGICTDLVANTSCGPCGPFMGRCQYKCSGHLGIGGITSNOQTASSNH 180
DB	121 NECEAEPCRNIGICTDLVANTSCGPCGPFMGRCQYKCSGHLGIGGITSNOQTASSNH 180
QY	181 RALFGLQKWPYYAALNKKGLINMATAENDRW-PMIQTV 220
DB	181 RALFGLQKWPYYAALNKKGLINMATAENDRW-PMIQTV 221
RESULT 7	
W10370	
ID	W10370 standard; Protein; 102 AA.
XX	
AC	W10370;
XX	
DT	03-MAY-1997 (first entry)
XX	
DE	Human Del-1 splice variant partial sequence.
XX	
KM	Del-1; developmentally-regulated endothelial cell locus-1;
KM	signal transduction; cancer; tumour marker; angiogenesis;
KM	diabetic retinopathy; rheumatoid arthritis; endometriosis;
KM	cardiac ischemia; stroke; vascular disease; wound healing;
KM	vulnerary; bone formation; diagnosis; therapy.
OS	
OS	Homo sapiens.
XX	
FH	
FT	Key Location/Qualifiers
FT	Domain 3..36
FT	/label= EGF-1
FT	/note= "epidermal growth factor-like domain 1"
FT	Domain 45..83
FT	/label= EGF-2
FT	/note= "epidermal growth factor-like domain 2"
FT	Domain 90..102
FT	/label= EGF-3

FT	/note= "N-terminal portion of epidermal growth factor-like domain 3"
FT	
XX	
PN	W096640769-A1.
XX	
PD	19-DEC-1996.
XX	
PE	05-JUN-1996; 96WO-US09456.
XX	
PR	07-JUN-1995; 95US-0480229.
XX	
PA	(PROG-) PROGENITOR INC. (UYVA-) UNIV VANDERBILT.
XX	
PI	Hogan B, Quertermous T, Snodgrass HR, Zupancic TJ;
XX	
DR	WPJ; 1997-052233/05.
XX	
PT	New developmentally regulated endothelial cell locus-1 (del-1) gene
PT	- used to develop prods. for the diagnosis and treatment of cancer
PT	and conditions involving abnormal angiogenesis
XX	
PS	Example; Page 91; 137pp; English.
CC	A polypeptide sequence (M10370) shows a portion of a variant form
CC	of human developmentally-regulated endothelial cell locus-1 (del-1)
CC	polypeptide. In comparison with the major form (M10365) of human
CC	Del-1, 10 amino acid residues are missing between EGF-like domains
CC	1 and 2 of the Del-1 variant. This is a result of alternative
CC	splicing (see also J47340).
XX	
SQ	Sequence 102 AA:
Query Match	42.9%; Score 542; DB 18; Length 102;
Best Local Similarity	82.1%; Pred. No. 7, 2e-29;
Matches 92; Conservative 3; Mismatches 7; Indels 10; Gaps	
OY	24 DICNPNCENGICISGLADSFSCPCPGFAPGNCSVVEVASDEKRTSAGPCIPNC 83 1 dlcidpncengicilpglavgsfscpcgpfcdpnssvvev-----gpcldpnc 50
DQ	84 HNGGTCEISEARGMFFIGYCKCPRGFGHCHOHNINCEAEPPCNGICT 135 51 hngtceiseayrgdflfigyckprgringlqchminceavepcknngict 102
RESULT 8	
Y94454	
ID	Y94454 standard; Protein; 426 AA.
AC	Y94454;
XX	
DT	11-SEP-2000 (first entry)
XX	
DE	Mouse lactadherin protein.
XX	
KW	Human; lactadherin; MGF-E8; anti-tumour; immune response; exosome; dendritic cell.
XX	
OS	Mus sp.
XX	
FH	Key Location/Qualifiers
FT	Peptide 1..22 /label= Secretion_signal
FT	Protein 23..426 /label= Lactadherin
FT	Binding-site 87..89 /label= Integrin_binding_site
FT	Misc-difference 93..111 /note= "encoded by Gr"
XX	
PN	EPI004664-A1.

```

XX 31-MAY-2000.
PD
XX
XX 24-NOV-1998; 98EP-0402925.
PF
XX 24-NOV-1998; 98EP-0402925.
PR
XX
XX (INRM ) INSEPM INST NAT SANTE & RECH MEDICALE.
PA (CURT-) INST CURIE.
XX
XX WPI: 2000-352597/31.
DR N-PSDB; A27141.
XX
XX Chimeric isolated (human) lactadherin polypeptide that functions as an
PT adaptor of cross-priming to eliminate pathogenic antigens, e.g. in
PT tumours -
XX
XX Disclosure; Page 12; 20pp; English.
XX
XX Lactadherin protein was found in exosomes produced by dendritic cells.
CC The protein is involved in the phagocytosis of particulate antigens by
CC dendritic cells. Exosomes produced by dendritic cells exposed to
CC tumour antigens induce potent immune responses. Lactadherin or variants
CC of it may be used in the mediation of an immune response. Variants of
CC lactadherin may be used for inhibition and/or stimulation of the
CC cross-priming of antigens and stimulation of the phagocytosis of
CC antigens by dendritic cells. Compositions derived from lactadherin can
CC also be used to monitor an immune response, more specifically a CTL
CC (cytotoxic T-lymphocyte) response and also to produce CTLs specific
CC for a selected antigen. The present sequence is the mouse
CC lactadherin protein.
XX
XX Sequence 426 AA;
SQ
Query Match 29.38; Score 369.5; DB 21; Length 426;
Best Local Similarity 36.24; Pred. No. 4.1e-17;
Matches 72; Conservative 26; Mismatches 50; Indels 51; Gaps 4;
QY 23 GDICNPNCENGICGLADDSFSCGCEGFAFGPNCSVVEVADDEKRTSAGPCIPNP 82
DB 25 gdfcdsalcngtclctg-gqndiylcpepfctgvcne-----tergpcspnp 72
QY 83 CHNGSTCEIS-EAYRGDTFTIGYCKPCRGNGIHCOHINECEAPCRNGICTDLVANY 141
DB 73 cyndakclvltldtgrgdlftfeyicgpcvysgsh----- 106
QY 142 SCECPGEFEMRNCQYKCSGHLGIBGIIISNQITASSNHRALFGLQKWPYVALNKKGL 201
DB 107 -----cectscqlgmeqgaladsgiasyvyngfmqlgrwqpeparlyrtgl 154
QY 202 INAWTAENDRMPWIOYTV 220
DB 155 vnaehasnyskpwigynl 173
RESULT 9
ID Y06816 standard; protein; 2471 AA.
XX
XX Y06816;
AC
XX
XX 05-JUL-1999 (first entry)
DT
XX
XX Human Notch2 (humN2) protein sequence.
XX
XX Notch: Nec; Ntm; epidermal growth factor; EGF; repeat domain; cancer;
KM transmembrane domain; cell-fate disorder; proliferative disease;
XX signal transduction; human; Notch2; humN2.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH

```

```

FT Cleavage-site 25
FT /note="potential signal cleavage site"
FT Region 26..1413
FT /note="EGF repeat region"
FT Region 1425..1537
FT /note="Lin-12/notch repeats"
FT Domain 1678..1699
FT /note="transmembrane domain"
FT Region 1822..2035
FT /note="Ankyrin repeats"
FT Region 2130..2133
FT /note="nuclear localisation signal (NLS) sequence"
FT Modified-site 2149..2152
FT /note="putative CKII phosphorylation site"
FT Modified-site 2156..2159
FT /note="putative cdc2 phosphorylation site"
FT Modified-site 2166..2169
FT /note="putative cdc2 phosphorylation site"
FT Region 2172..2174
FT /note="nuclear localisation signal (BNRS) sequence"
FT Region 2385..2445
FT /note="PEST-containing region"
FT
FT MO9904746-A2.
FT
FT 04-FEB-1999.
FT
FT 23-JUL-1998; 98WO-US15333.
FT
FT 23-JUL-1997; 97US-0899232.
FT
FT (UYRA ) UNIV YALE.
FT
FT Artavanis-Tsakonas S, Qi H, Rand MD;
FT WPI: 1999-153290/13.
FT
FT Measuring and detecting activation of Notch protein in cells from
PT cell surface expression - useful for, e.g. identifying modulators of
PT activation, potentially useful for treating cancer
XX
XX Disclosure; Flg 2A-D; 94pp; English.
XX
XX The invention relates to methods of measuring activation of Notch in a
XX cell. Activated notch can be detected from: (1) expression of: (a) Notch
XX on the cell surface, or (b) one or both of the Notch cleavage products
XX Nec and Ntm; (11) presence of Notch fragments, i.e. one or both of an
XX N-terminal fragment (F1) terminating between the epidermal growth factor
XX -like (EGF) repeat domain and the transmembrane domain or a C-terminal
XX fragment (F2) that has its N-terminus between the specified domains, or
XX fragments of molecular weights about 270, 200, 170, 140, 110, 100, 90 and
XX 85 kDa; or (111) presence of a Notch heterodimer containing a linkage
XX sensitive to reducing agents. The method is used to study or manipulate
XX differentiation processes and to screen/diagnose cell-fate disorders
XX (particularly cancer or other proliferative diseases involving abnormal
XX Notch activation). Modulators of Notch activation are useful for studying
XX or manipulating differentiation and possibly for treating cancer. The
XX F1-F2 heterodimer is the active form of Notch that mediates signal
XX transduction and binds ligands such as Delta and Serrate. F1 and F2 are
XX generated by proteolysis of full-length Notch in the trans-Golgi. The
XX present sequence represents the human Notch2 (humN2) sequence.
XX
XX Sequence 2471 AA;
SQ
Query Match 25.28; Score 318; DB 20; Length 2471;
Best Local Similarity 41.18; Pred. No. 4.1e-13;
Matches 58; Conservative 15; Mismatches 44; Indels 24; Gaps 4;
QY 24 DICNPNCENGICGLADDSFSCGCEGFAFGPNCSVVEVADDEKRTSAGPCIPNP 83
DB 913 ddcianpcngsgscmdyv--ntfscclpftgdkc-----tdmncisepc 958

```


CC	C-delta-1 polypeptide (W11719) is the chick homologue of Drosophila
CC	Delta, a protein that binds to Notch protein. Expression of
CC	C-delta-1 correlates with onset of neurogenesis. The C-delta-1
CC	amino acid sequence was deduced from a cDNA clone (T58897) obtd.
CC	from chick stage 4-6 embryos. An alternatively spliced variant
CC	(W00876) was also isolated, and mouse (W11720) and human (W11721-
CC	38) Delta-1 polypeptides have been identified. Delta-1 proteins
CC	can be used to treat or prevent disorders characterised by
CC	increased Notch activity, such as cervical, breast, lung or colon
CC	cancer, melanoma or seminoma, and nervous system disorders or to
CC	promote tissue regeneration and repair.
SQ	Sequence 727 AA:
Query Match	24.9%; Score 314.5; DB 18; Length 727;
Best Local Similarity	29.1%, Pred.No. 2,5e-13;
Matches 72; Conservative 24; Mismatches 74; Indels 77; Gaps	77
OY	26 CNPNPENGIGICISGLADSFSCCECFEGFAGNCSSVVEASDEEKPPTSAGCIPNPCHN 85
DB	339 clanpekngsgstd--lensysctcoppgyfygnc-----lsatmcadgpcfn 384
OY	86 GG-----TCETSEAYRG-----DTFTGYCKCPKPFNG 113
DB	385 ggcrlndpgygsccrplpygsfneckkidycssspcaagacvdlgnsgylcqcgafteq 444
OY	114 HGHQHINCEAEPCRNNGGICTDLVANYSCCEPEFGMRNCVKCSG-----HLGLEGG 167
DB	445 rhdddnvdcastpcvnvgtlcqgvnydysetcpeppynghkncstprvsrchnpchgatch 504
OY	168 IISNQITASSNNHRLFLGLQKWYPYYAALNKGLI-----NAWTAENDRMP 214
DB	505 esnryvccearg-----yglnqcflilpeppgpvlavdftekylegnsqtfp 552
OY	215 WIQOYWVG 221
DB	553 wlvacag 559
RESULT 11	
ID Y79029	standard; protein: 728 AA.
AC Y79029;	
DT 06-JUN-2000	(first entry)
XX DE	Chick delta protein amino acid sequence.
XX KW	Delta; Kuz; Notch; differentiation; cervical; breast; colon; cancer;
KW lung; melanoma; seminoma; central nervous system disorder; psoriasis;	
KW tissue regeneration; liver cirrhosis; keloid formation; baldness;	
KW inner ear disorder; chick.	
OS Gallus sp.	
XX WO200002897-A2.	
PD 20-JAN-2000.	
XX PF	13-JUL-1999; 99WC-US15817.
XX PR	13-JUL-1998; 98US-0092513.
XX PA	19-OCT-1998; 98US-0104834.
PA (UYVA) UNIV YALE.	
PI Artavanis-Tsakonas S, Rand MD, Qi H;	
DR WPI; 2000-282852/24.	
XX	

PT New cleavage peptide, nucleic acids and antibodies useful for
PT diagnosis, prevention and treatment of cancer, disorders of central
PT nervous system, cirrhosis and psoriasis
PS Claim 1; Fig 3; 177pp; English.

XX This sequence represents the chick delta protein amino acid sequence.
CC Delta is a topotypic protein that contains a sequence which is cleaved
CC by the metalloprotease-disintegrin Kuzbanian (kuz). Cleavage by kuz
CC results in two fragments, a soluble amino terminal fragment consisting
CC essentially of the extracellular domain, and a membrane bound fragment
CC consisting of the transmembrane domain and the intracellular domain. The
CC soluble fragment is able to bind to Notch. Delta plays a key role in
CC differentiation, and therefore detection and measurement of delta
CC activation is important in the study of differentiation. The invention
CC relates to the delta cleavage peptides (the active fragment), and to
CC methods for detecting and measuring delta activation. Delta cleavage
CC peptides, and chimeric proteins are useful for modulating the activity of
CC Notch, delta or kuz or at least one of the signalling pathways in a cell
CC or organism, expressing Notch. By contacting a cell with kuz protein or
CC nucleic acid or its antibody, the activity or levels of delta protein is
CC modulated and vice versa. A delta cleavage peptide or its derivative
CC is capable of binding kuz protein is useful for treating or preventing a
CC disease or disorder associated with increased delta activity or
CC expression such as cervical, breast, colon or lung cancer, melanoma or
CC seminoma in humans. A recombinant cell comprising a delta peptide is
CC useful for treating or preventing central nervous system disorders. A
CC delta cleavage peptide is useful for the diagnosis of diseases or
CC disorders associated with increased levels of Notch-delta cleavage
CC binding activity comprising measuring the ability of Notch-delta cleavage
CC peptides in a sample to bind kuz protein. A complex of delta cleavage
CC kuz is useful for diagnosing or screening for the presence of, or
CC predisposition to developing a disease or disorder associated with
CC aberrant levels of the complex, comprising measuring the level of
CC functional activity of the complex or RNA encoding delta or kuz in a
CC sample. The delta cleavage peptide is also useful for promoting tissue
CC regeneration and repair, for treating liver cirrhosis, keloid formation,
CC psoriasis, baldness and degenerative or traumatic disorders of the
CC sensory epithelium of the inner ear.

SQ Sequence 728 AA;

Query Match 24.9%; Score 314.5; DB 21; Length 728;
Best Local Similarity 29.1%; Pred. No. 2.5e-13;
Matches 72; Conservative 24; Mismatches 74; Indels 77; Gaps 7;

QY 26 CNPNCENGICLSGLADSFSCPEGFAGNCSSVVEASDEKPTISAGPCLIPNCHN 85
DB 339 cdampcknggsectd--lensysctcpptfygknc-----lsamtcadgpcfn 384
QY 86 GG-----TCEISEAYRG-----DTFGYCKCPRGNG 113
DB 385 ggrctcdnpdgysrcrclpygfsfneckidysscspcngagcdlgnyslcqcaqftg 444
QY 114 IHCQHNINCEAEPRNCGICTDLVANTSCPEGFEMRNCQYKCSG-----HLGTEGG 167
DB 445 rhcdnhyddcasfpcvngtcdgdyndscfcpynkncstpsrcehpnngatch 504
QY 168 IISNOQITASSNHRALFGLQKWPYALNKKGLI-----NMTAENDRWP 214
DB 505 ersnuyveecaag-----yglncqfllpeppgprvltvdfekykeqgnsqfip 552
QY 215 WIQVTVG 221
DB 553 wlvacag 559

RESULT 12
W00876
ID W00876 standard: Protein; 740 AA.
XX
AC W00876; *

XX 28-APR-1997 (first entry)
DT
XX
XX C-Delta-1 polypeptide (alternatively spliced variant).
DE
XX
XX C-Delta-1: cell proliferation; nervous system disorder;
KW tissue regeneration; Notch; cervix cancer; breast cancer;
KW lung cancer; colon cancer; melanoma; seminoma;
XX neurogenesis; therapy.
OS
XX Gallus sp.
FH
FT Key
FT Domain
FT 184..228
FT /label= DSL
FT 229..261
FT /label= EGF1
FT 262..292
FT /label= EGF2
FT 293..332
FT /label= EGF3
FT 333..370
FT /label= EGF4
FT 371..409
FT /label= EGF5
FT 410..447
FT /label= EGF6
FT 448..485
FT /label= EGF7
FT 486..523
FT /label= EGF8
FT 524..534
FT /label= EGF9
FT 535..579
FT /label= TM
FT /note= "transmembrane domain"
FT
FT
XX
XX W09701571-AI.
PD 16-JAN-1997.
XX
XX 28-JUN-1996; 96WO-US11178.
XX
XX 28-JUN-1995; 95US-0000589.
XX
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX (UYTA) UNIT YALE.
XX
XX Artavanis-Tsakonas S, Gray GE, Henrique D, Ish-Horowicz D;
PI Lewis J;
XX
XX WPI: 1997-100159/09.
XX N-PSDB; T58898.
XX
XX New vertebrate Delta protein, DNA and antibodies - for treating and
PT preventing cancer, nervous system disorders and for tissue
PT regeneration
PT
PT
XX
PS Disclosure: Fig 2; 135pp; English.
XX
XX C-delta-1 polypeptide (W00876) is the chick homologue of Drosophila
CC Delta, a protein that binds to Notch protein. Expression of
CC C-delta-1 correlates with onset of neurogenesis. The C-delta-1
CC amino acid sequence was deduced from a cDNA clone (T58898) obtd.
CC from chick stage 4-6 embryos. A shorter version (W58877) of
CC C-delta-1, lacking the 12 C-terminal amino acids of the longer
CC version, was also isolated, and mouse (W11720) and human (W11721-
CC 38) Delta-1 polypeptides have been identified. Delta-1 proteins
CC can be used to treat or prevent disorders characterised by
CC increased Notch activity, such as cervical, breast, lung or colon
CC cancer, melanoma or seminoma, and nervous system disorders or to
CC promote tissue regeneration and repair.
XX

PD 29-MAY-1997.
XX
XX 15-NOV-1996; 96MO-JP03356.
XX
XX 30-NOV-1995; 95JP-0311811.
PR 17-NOV-1995; 95JP-0299611.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
XX
XX Itoh A. Sakano S;
XX
XX WPI: 1997-298110/27.
XX
XX Peptide(s) encoded by human genes delta-1 and serrate-1 - suppress
PT proliferation and differentiation of undifferentiated human blood
PR cells
XX
XX Claim 4; Page 61-64; 114pp; Japanese.
XX
XX The present sequence represents a polypeptide which suppresses
CC proliferation and differentiation of undifferentiated cells such
CC as neurons and blood cells. The polypeptide may be used for the
CC prevention and control of disorders involving undifferentiated
CC cells, such as leukaemia and malignant tumours, and improvement of
CC blood formation, e.g. after immunosuppression.
XX
SQ Sequence 702 AA;

Query Match 24.1%; Score 304; DB 18; Length 702;
Best Local Similarity 28.3%; Pred. No. 1.2e-12;
Matches 70; Conservative 25; Mismatches 78; Indels 74; Gaps 8;

QY 24 DICNPNCENGICLSGLADDSFSCPEGFAGPNCSSVEVASEDEKPTSGAGCIPNPC 83
DB 309 decdpckngsgctd--lensysctcpfpfygkice-----lsamtcadgpc 354
QY 84 HNGGTCEISE-----AYRG-----DTEIGYVCKCPRGF 111
DB 355 fnggcsdspdggysrcrcpvygsgfncckkidyccsspsngakcvdldgdaylcrqagf 414
QY 112 NGHCQHINIECEAEPCRNCGICTDLVANYSCCEGCEFGMRNCOTKCS--GHGIEGGIT 169
DB 415 sgrhcdnvdvdcasspcangtctdgvndfscctppgytgrncsapvrcehapchngat 474
QY 170 SNOQITASSNHRALFGLOKMY-----PYAALNKKGLINAWTAENDRWP 214
DB 475 cher-----ghryvcecargyggnpcqfllpelppgavvdltek-----legggpfp 523
QY 215 WIOYTVG 221
DB 524 wvavcag 530

RESULT 15
W75495
ID W75495 standard; Protein; 702 AA.
XX
AC W75495;
XX
XX 27-APR-1999 (first entry)
XX
XX Truncated human delta-1 protein #3.
XX
XX Human; delta-1 protein; ligand; notch; drug; vascular cell; primer; PCR;
KW amplification; truncation.
XX
XX Homo sapiens.
XX
XX JP10316582-A.
XX
XX 02-DEC-1998.
XX

PF 14-MAY-1997; 97JP-0124062.
XX
XX 14-MAY-1997; 97JP-0124062.
XX
XX (ASAH) ASAH KASEI KOGYO KK.
XX
XX WPI: 1999-076401/07.
XX
XX Vascular cell controlling agent comprises polypeptide - which is
PT human notch ligand and is used as drug
XX
XX Claim 2; Page 14-16; 21pp; Japanese.
XX
XX This sequence represents a truncated human delta-1 protein corresponding
CC to amino acids 1-702 of the mature protein (see W75492). The delta-1
CC protein is a ligand of the human notch protein and the protein or
CC fragments, especially W75493-W75495, can be used as a drug to control
CC vascular cells. The sequences were isolated and the truncated fragments
CC were generated using the primers X16818-X16831.
XX
SQ Sequence 702 AA;

Query Match 24.1%; Score 304; DB 20; Length 702;
Best Local Similarity 28.3%; Pred. No. 1.2e-12;
Matches 70; Conservative 25; Mismatches 78; Indels 74; Gaps 8;

QY 24 DICNPNCENGICLSGLADDSFSCPEGFAGPNCSSVEVASEDEKPTSGAGCIPNPC 83
DB 309 decdpckngsgctd--lensysctcpfpfygkice-----lsamtcadgpc 354
QY 84 HNGGTCEISE-----AYRG-----DTEIGYVCKCPRGF 111
DB 355 fnggcsdspdggysrcrcpvygsgfncckkidyccsspsngakcvdldgdaylcrqagf 414
QY 112 NGHCQHINIECEAEPCRNCGICTDLVANYSCCEGCEFGMRNCOTKCS--GHGIEGGIT 169
DB 415 sgrhcdnvdvdcasspcangtctdgvndfscctppgytgrncsapvrcehapchngat 474
QY 170 SNOQITASSNHRALFGLOKMY-----PYAALNKKGLINAWTAENDRWP 214
DB 475 cher-----ghryvcecargyggnpcqfllpelppgavvdltek-----legggpfp 523
QY 215 WIOYTVG 221
DB 524 wvavcag 530

Search completed: May 23, 2001, 06:15:18
Job time: 232 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:18:15 ; Search time 179.67 Seconds
(without alignments)
197.856 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 1262
Sequence: 1 MKHLVAMLVGLSLGVPQF.....INMTAENDRWPWQVTVG 221

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 segs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCRUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1262	100.0	221	16	US-09-237-981-29
2	1242	98.4	480	8	US-08-480-229B-10
3	1242	98.4	480	10	US-08-659-235B-10
4	1242	98.4	480	16	US-09-237-981-10
5	1133.5	89.8	481	10	US-08-659-235B-29
6	1133.5	89.8	513	10	US-08-480-229B-14
7	1133.5	89.8	513	8	US-08-659-235B-14
8	1133.5	89.8	513	16	US-09-237-981-14
9	1007	79.8	203	23	US-60-233-644-77
10	995	78.8	185	23	US-60-230-435-1806

11	735.5	58.3	449	23	US-60-207-315-404	Sequence 404, App
12	533	42.2	362	23	US-60-207-315-523	Sequence 523, App
13	532	42.2	311	23	US-60-230-435-1519	Sequence 1519, App
14	389	30.8	463	19	US-09-582-340-4	Sequence 4, Appl
15	324.5	25.7	2523	1	PCT-US99-15817-3	Sequence 3, Appl
16	324.5	25.7	2523	12	US-08-899-232-3	Sequence 3, Appl
17	324.5	25.7	2523	15	US-09-121-457-3	Sequence 3, Appl
18	319	25.3	62	23	US-60-192-739-3075	Sequence 3075, App
19	318	25.2	1112	23	US-60-229-515-1196	Sequence 1196, App
20	318	25.2	1361	23	US-60-230-445-1730	Sequence 1730, App
21	318	25.2	2090	23	US-60-213-177-839	Sequence 839, App
22	318	25.2	2146	23	US-60-213-177-839	Sequence 473, App
23	318	25.2	2471	1	PCT-US99-15817-1	Sequence 1, Appl
24	318	25.2	2471	4	US-08-083-590-19	Sequence 19, Appl
25	318	25.2	2471	12	US-08-899-232-1	Sequence 1, Appl
26	318	25.2	2471	13	US-08-937-132-19	Sequence 19, Appl
27	318	25.2	2471	13	US-08-947-956-19	Sequence 19, Appl
28	318	25.2	2471	15	US-09-121-457-1	Sequence 1, Appl
29	317.5	25.2	721	1	PCT-US99-15817-6	Sequence 6, Appl
30	317.5	25.2	721	12	US-08-832-633-16	Sequence 16, Appl
31	317.5	25.2	721	13	US-08-832-633-5	Sequence 5, Appl
32	317.5	25.2	721	18	US-09-483-673-7	Sequence 7, Appl
33	317.5	25.2	721	18	US-09-483-673-7	Sequence 7, Appl
34	314.5	24.9	728	13	US-08-981-392-2	Sequence 13, Appl
35	314.5	24.9	728	13	US-08-981-392-2	Sequence 13, Appl
36	314	24.9	729	18	US-08-832-633-17	Sequence 17, Appl
37	314	24.9	729	18	US-09-483-673-8	Sequence 8, Appl
38	314	24.9	729	18	US-09-483-674-8	Sequence 8, Appl
39	308.5	24.4	2601	23	US-60-171-625-153	Sequence 153, App
40	308.5	24.4	2601	23	US-60-173-464-6157	Sequence 6157, App
41	308.5	24.4	2634	23	US-60-191-637-7611	Sequence 7611, App
42	308.5	24.4	2634	23	US-60-191-681-5934	Sequence 5934, App
43	305.5	24.2	713	12	US-08-832-633-14	Sequence 14, Appl
44	305.5	24.2	713	18	US-09-483-673-5	Sequence 5, Appl
45	305.5	24.2	713	18	US-09-483-674-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-237-981-29
Sequence 981, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quereternous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 221 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-237-981-29

Query Match 100.0% Score 1262: DB 16: Length 221:
Best Local Similarity 100.0% Pred. No. 5.1e-104:
Matches 221: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 MKHLVAAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
DB 1 MKHLVAAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
QY 61 SYVEVASDEKPTPSAGPCIPNCPCHNGTCEISAYRGDTFTGYVCKCPGFGNGIHCOHNI 120
DB 61 SYVEVASDEKPTPSAGPCIPNCPCHNGTCEISAYRGDTFTGYVCKCPGFGNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
QY 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTVG 221
DB 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTVG 221

RESULT 2
US-08-480-229B-10
Sequence 10, Application US/08480229B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229B-10

Query Match 98.4% Score 1242: DB 8: Length 480:
Best Local Similarity 98.2% Pred. No. 7.2e-102:
Matches 216: Conservative 2: Mismatches 2: Indels 0: Gaps 0:

QY 1 MKHLVAAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
DB 1 MKHLVAAWLLVGLSLGVPOFGKGDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
QY 61 SYVEVASDEKPTPSAGPCIPNCPCHNGTCEISAYRGDTFTGYVCKCPGFGNGIHCOHNI 120
DB 61 SYVEVASDEKPTPSAGPCIPNCPCHNGTCEISAYRGDTFTGYVCKCPGFGNGIHCOHNI 120
QY 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
QY 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTV 220
DB 181 RALFGLQKWYPYAAALNKKGLINAWTAENDRWPWIOQTV 220

RESULT 3
US-08-659-235B-10
Sequence 10, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Brigid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-10

Query Match 98.4%; Score 1242; DB 10; Length 480;
Best Local Similarity 98.2%; Pred. No. 7.2e-102;
Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKHLVAAWLLVGLSGVPOFGKGDICNPNCENGICLSGLADDSFSCPCPFGAGPNC 60
DB 1 MKHLVAAWLLVGLSGVPOFGKGDICNPNCENGICLSGLADDSFSCPCPFGAGPNC 60

QY 61 SVEVASDEEKPSTAGPCIPNPNCHNGTCSEAVRGDTFTGYVCKCRPGNGIHCOHNI 120
DB 61 SVEVASDEEKPSTAGPCIPNPNCHNGTCSEAVRGDTFTGYVCKCRPGNGIHCOHNI 120

QY 121 NECEAEPCRNNGICTDLVANYSCPCGPEFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCPCGPEFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180

QY 181 RALFGLQWYPPYALNKKGLINAWTAENDRWPWIOQTV 220
DB 181 RALFGLQWYPPYALNKKGLINAWTAENDRWPWIOQTV 220

RESULT 4
US-09-237-981-10
Sequence 10, Application US/09237981
GENERAL INFORMATION:
APPLICANT: Quartermours, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/237,981
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 480 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-237-981-10

Query Match 98.4%; Score 1242; DB 16; Length 480;
Best Local Similarity 98.2%; Pred. No. 7.2e-102;
Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKHLVAAWLLVGLSGVPOFGKGDICNPNCENGICLSGLADDSFSCPCPFGAGPNC 60

DB 1 MKHLVAAWLLVGLSGVPOFGKGDICNPNCENGICLSGLADDSFSCPCPFGAGPNC 60
QY 61 SVEVASDEEKPSTAGPCIPNPNCHNGTCSEAVRGDTFTGYVCKCRPGNGIHCOHNI 120
DB 61 SVEVASDEEKPSTAGPCIPNPNCHNGTCSEAVRGDTFTGYVCKCRPGNGIHCOHNI 120

QY 121 NECEAEPCRNNGICTDLVANYSCPCGPEFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180
DB 121 NECEAEPCRNNGICTDLVANYSCPCGPEFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180

QY 181 RALFGLQWYPPYALNKKGLINAWTAENDRWPWIOQTV 220
DB 181 RALFGLQWYPPYALNKKGLINAWTAENDRWPWIOQTV 220

RESULT 5
US-08-659-235B-29
Sequence 29, Application US/08659235B
GENERAL INFORMATION:
APPLICANT: Quartermours, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235B
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235B-29

Query Match 89.8%; Score 1133.5; DB 10; Length 481;
Best Local Similarity 90.0%; Pred. No. 3e-92;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 MKHLVAAWLLVGLSGVPOFGKGDICNPNCENGICLSGLADDSFSCPCPFGAGPNC 60
DB 1 MKRVAWVLLVGLSGVPOFGKGDICDPNPNCHNGTCSEAVRGDTFTGYVCKCRPGNGIHCOHNI 60

QY 61 SVEVASDEEKPSTAGPCIPNPNCHNGTCSEAVRGDTFTGYVCKCRPGNGIHCOHNI 120
DB 61 SVEVASDEEKPSTAGPCIPNPNCHNGTCSEAVRGDTFTGYVCKCRPGNGIHCOHNI 120

QY 121 NECEAEPCRNNGICTDLVANYSCPCGPEFMRNCQYKCSGHLGIEGGIISNOQITASSNH 180

Db 121 NECEVEPCCKNGICTDLVANYSCCEPGCEMGRNCQYKCSGFLGEGGITSNOQTASTH 180
QY 181 RALFGLQKWPYYAALNKKGLINMTAAENDRWP-WIQTV 220
Db 181 RALFGLQKWPYYAALNKKGLINMTAAENDRWMKRWIQINL 221

RESULT 6

US-08-480-229B-14
; Sequence 14, Application US/08480229B
; GENERAL INFORMATION:
; APPLICANT: Quartermours, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-229B-14

Query Match 89.8%; Score 1133.5; DB 8; Length 513;
Best Local Similarity 90.0%; Pred. No. 3.2e-92;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 MKHLVAAMLVLGLSLGVPOFGKGDICNPNCENGICISGLADSFSCCEPGFAGPNC 60
Db 33 MKRSVAVMLLVGLSLGVPOFGKGDICDPNCPENGICLPLGLAVSFSCECPDGTDPNCS 92
QY 61 SVEEVADEEKP7SAGPCIPNCHNGTCEISEAYRGDTF1GYVCKCPRGFNGIHCOHNI 120
Db 93 SVEEVADEEPT7SAGPCIPNCHNGTCEISEAYRGDTF1GYVCKCPRGFNGIHCOHNI 152
QY 121 NECEAPPCNGGICTDLVANYSCCEPGCEMGRNCQYKCSGHLGIEGGITSNOQTASSNH 180
Db 153 NECEVEPCCKNGICTDLVANYSCCEPGCEMGRNCQYKCSGFLGIEGGITSNOQTASTH 212
QY 181 RALFGLQKWPYYAALNKKGLINMTAAENDRWP-WIQTV 220
Db 213 RALFGLQKWPYYAALNKKGLINMTAAENDRWMKRWIQINL 253

RESULT 7

US-08-659-235B-14
; Sequence 14, Application US/08659235B
; GENERAL INFORMATION:
; APPLICANT: Quartermours, Thomas
; APPLICANT: Hogan, Bridgid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235B
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-659-235B-14

Query Match 89.8%; Score 1133.5; DB 10; Length 513;
Best Local Similarity 90.0%; Pred. No. 3.2e-92;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 MKHLVAAMLVLGLSLGVPOFGKGDICNPNCENGICISGLADSFSCCEPGFAGPNC 60
Db 33 MKRSVAVMLLVGLSLGVPOFGKGDICDPNCPENGICLPLGLAVSFSCECPDGTDPNCS 92
QY 61 SVEEVADEEKP7SAGPCIPNCHNGTCEISEAYRGDTF1GYVCKCPRGFNGIHCOHNI 120
Db 93 SVEEVADEEPT7SAGPCIPNCHNGTCEISEAYRGDTF1GYVCKCPRGFNGIHCOHNI 152
QY 121 NECEAPPCNGGICTDLVANYSCCEPGCEMGRNCQYKCSGHLGIEGGITSNOQTASSNH 180
Db 153 NECEVEPCCKNGICTDLVANYSCCEPGCEMGRNCQYKCSGFLGIEGGITSNOQTASTH 212
QY 181 RALFGLQKWPYYAALNKKGLINMTAAENDRWP-WIQTV 220
Db 213 RALFGLQKWPYYAALNKKGLINMTAAENDRWMKRWIQINL 253

RESULT 8

US-09-237-981-14
; Sequence 14, Application US/09237981
; GENERAL INFORMATION:
; APPLICANT: Quartermours, Thomas
; APPLICANT: Hogan, Bridgid


```

; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/237,981
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235
; FILING DATE: 05-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-237-981-14

Query Match          89.8%; Score 1133.5; DB 16; Length 513;
Best Local Similarity 90.0%; Pred. No. 3.2e-92;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 MKHLVAAMLVGLSLGVPOFGKSDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 60
   |||||||
DB 33 MKRSVAVMLLVGLSLGVPOFGKSDICNPNCENGICLSGLADDSFSCCEPGEFAGPNC 92
QY 61 SYVEVASDEKPTISAGPCINPCNHNGTCEISBAYRGDFFIGYVCKCPRGFNGIHCOHNT 120
   |||||||
DB 93 SYVEVASDEKPTISAGPCINPCNHNGTCEISBAYRGDFFIGYVCKCPRGFNGIHCOHNT 152
QY 121 NECEAEPCRRNGICTDLVANYSCCEPGEFMRNCOYKCSHLCIEGIIISNOQITASSNH 180
   |||||||
DB 153 NECEAEPCRRNGICTDLVANYSCCEPGEFMRNCOYKCSHLCIEGIIISNOQITASSNH 212
QY 181 RALFGLQKWPYAAIINKKGLINAMTAENDRMP-WIQVTV 220
   |||||||
DB 213 RALFGLQKWPYAAIINKKGLINAMTAENDRMP-WIQVTV 253

RESULT 9
US-60-233-644-77
; Sequence 77, Application US/60233644
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
; TITLE OF INVENTION: PROTEIN, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: CLO00824
; CURRENT APPLICATION NUMBER: US/60/233,644

```

```

; CURRENT FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 203
; TYPE: PRT
; ORGANISM: HUMAN
; US-60-233-644-77

Query Match          79.8%; Score 1007; DB 23; Length 203;
Best Local Similarity 88.4%; Pred. No. 1.9e-81;
Matches 175; Conservative 3; Mismatches 10; Indels 10; Gaps 1;

QY 24 DICNPNCENGICLSGLADDSFSCCEPGEFMRNCOYKCSHLCIEGIIISNOQITASSNH 83
   |||||||
DB 12 DICNPNCENGICLSGLADDSFSCCEPGEFMRNCOYKCSHLCIEGIIISNOQITASSNH 61
QY 84 HNGTCEISEAYRGDTFTIGYVCKCPRGFNGIHCOHINCEAEPCRRNGICTDLVANYSC 143
   |||||||
DB 62 HNGTCEISEAYRGDTFTIGYVCKCPRGFNGIHCOHINCEAEPCRRNGICTDLVANYSC 121
QY 144 ECPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNHRAFLQKWPYAAIINKKGLIN 203
   |||||||
DB 122 ECPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNHRAFLQKWPYAAIINKKGLIN 181
QY 204 AMTAENDRMPWIIQVTVG 221
   |||||||
DB 182 AMTAENDRMPWIIQVTVG 199

RESULT 10
US-60-230-435-1806
; Sequence 1806, Application US/60230435
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO00768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1806
; LENGTH: 185
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(185)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-60-230-435-1806

Query Match          78.8%; Score 995; DB 23; Length 185;
Best Local Similarity 88.7%; Pred. No. 1.9e-80;
Matches 172; Conservative 3; Mismatches 9; Indels 10; Gaps 1;

QY 24 DICNPNCENGICLSGLADDSFSCCEPGEFMRNCOYKCSHLCIEGIIISNOQITASSNH 83
   |||||||
DB 2 DICNPNCENGICLSGLADDSFSCCEPGEFMRNCOYKCSHLCIEGIIISNOQITASSNH 51
QY 84 HNGTCEISEAYRGDTFTIGYVCKCPRGFNGIHCOHINCEAEPCRRNGICTDLVANYSC 143
   |||||||
DB 52 HNGTCEISEAYRGDTFTIGYVCKCPRGFNGIHCOHINCEAEPCRRNGICTDLVANYSC 111
QY 144 ECPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNHRAFLQKWPYAAIINKKGLIN 203
   |||||||
DB 112 ECPGEFMRNCOYKCSGHLGIEGIIISNOQITASSNHRAFLQKWPYAAIINKKGLIN 171
QY 204 AMTAENDRMPWIIQ 217
   |||||||

```

Db 172 AWTAAENDRWPMIQ 185

RESULT 11

US-60-207-315-404
; Sequence 404, Application US/60207315

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; FILE REFERENCE: CL000601

; CURRENT APPLICATION NUMBER: US/60/207,315

; CURRENT FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 528

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 404

; LENGTH: 449

; TYPE: PRT

; ORGANISM: HUMAN

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(449)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-60-207-315-404

Query Match

Best Local Similarity 58.3%; Score 735.5; DB 23; Length 449;
Matches 135; Conservative 1; Mismatches 5; Indels 95; Gaps 1;

QY 77 PCIRPCHNGTCEISAFYGDPIFYVCKCPGFNGIHCH----- 118

Db 2 PCTPNCHNGTCEISAFYGDPIFYVCKCPGFNGIHCHCPALADPLSVPG 61

QY 119 ----- 118

Db 62 KPLPLSLIDNAPRRSSKNAPRISTSSGRKYPILSTLREPSDWEHVGPKAAMNTGLAMPVL 121

QY 119 -----NINECEAPCRNGIGICTDLVANYSCCEPGFPMGRNCQYKCSGH 161

Db 122 RGGILGQLLRPECGAPNINECEVEPCCKNGIGICTDLVANYSCCEPGFPMGRNCQYKCSGP 181

QY 162 LGIEGGIISNOQITASSNHRALFGLOKWPYVALNKKGLINAWTAENDRWPMIQ 217

Db 182 LGIEGGIISNOQITASSNHRALFGLOKWPYVALNKKGLINAWTAENDRWPMIQ 237

RESULT 12

US-60-207-315-523
; Sequence 523, Application US/60207315

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; FILE REFERENCE: CL000601

; CURRENT APPLICATION NUMBER: US/60/207,315

; CURRENT FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 528

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 523

; LENGTH: 362

; TYPE: PRT

; ORGANISM: HUMAN

US-60-207-315-523

Query Match

Best Local Similarity 42.2%; Score 533; DB 23; Length 362;
Matches 93; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 119 NINECEAPCRNGIGICTDLVANYSCCEPGFPMGRNCQYKCSGHIGIGIISNOQITASS 178

Db 52 DINECEVEPCCKNGIGICTDLVANYSCCEPGFPMGRNCQYKCSGPLGIGIGIISNOQITASS 111

QY 179 NHRALFGLOKWPYVALNKKGLINAWTAENDRWPMIQ 217

Db 112 THRALFGLOKWPYVALNKKGLINAWTAENDRWPMIQ 150

RESULT 13

US-60-230-435-1519

; Sequence 1519, Application US/60230435

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

; FILE REFERENCE: CL000768

; CURRENT APPLICATION NUMBER: US/60/230,435

; CURRENT FILING DATE: 2000-09-06

; NUMBER OF SEQ ID NOS: 2991

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1519

; LENGTH: 311

; TYPE: PRT

; ORGANISM: HUMAN

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(311)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-60-230-435-1519

Query Match 42.2%; Score 532; DB 23; Length 311;

Best Local Similarity 94.9%; Pred. No. 3.8e-39;
Matches 93; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 120 INCEAPCRNGIGICTDLVANYSCCEPGFPMGRNCQYKCSGHIGIGIISNOQITASSN 179

Db 2 INCEVEPCCKNGIGICTDLVANYSCCEPGFPMGRNCQYKCSGPLGIGIGIISNOQITASS 61

QY 180 HRALFGLOKWPYVALNKKGLINAWTAENDRWPMIQ 217

Db 62 HRALFGLOKWPYVALNKKGLINAWTAENDRWPMIQ 99

RESULT 14

US-09-582-340-4
; Sequence 4, Application US/09582340

; GENERAL INFORMATION:

; APPLICANT: INSERM

; APPLICANT: INSTITUT CURIE

; APPLICANT: CNRS

; TITLE OF INVENTION: Compositions and Methods using Lactadherin or Variants

; TITLE OF INVENTION: thereof

; FILE REFERENCE: Lactadherin

; CURRENT APPLICATION NUMBER: US/09/582,340

; CURRENT FILING DATE: 2000-09-15

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 463

; TYPE: PRT

; ORGANISM: mammalian

US-09-582-340-4

Query Match 30.8%; Score 389; DB 19; Length 463;

Best Local Similarity 38.8%; Pred. No. 2.8e-26;
Matches 78; Conservative 30; Mismatches 75; Indels 18; Gaps 5;

QY 23 GDICNPENCGIGICISGLADDSFSCCEPGFPMGRNCQYKCSGVYEVASDEKPTISAGCINP 82

Db 25 GDFCDSSILCNGTCTLTG-QDMNDIYCLCPGFTGLVCNE-----TERGFCSPNP 72

Best Local Similarity 36.1%; Pred. No. 1.3e-15;
Matches 48; Conservative 18; Mismatches 47; Indels 20; Gaps 6;

QY 25 IC-NPNCENGICLSGLADDSFSCCEPGEFAGPNCSSVVEASDEEKPPTSAGPC--IPN 81
Db 94 ICTQSPQNGGQCMYD--GGGEYHCVCILPFGHGRDC-----ERK---AGPCEQAGS 140
QY 82 PCHNGGTCISEAYRGDTFTIGYVCKCPRGFNGIHCOHININECEAPCRNGICTDLVANY 141
Db 141 PCRRGGCCODDGGFA---LNFCTCRCLVGFVGARCEVNVDDCLMRPCANGATCILDGINRF 196
QY 142 SCECPGEFMRNC 154
Db 197 SCLCPGEFAGRNC 209

RESULT 6
US-09-833-245-770

; Sequence 770, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 770
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-770

Query Match 19.1%; Score 241; DB 5; Length 383;
Best Local Similarity 36.1%; Pred. No. 1.3e-15;

Matches 48; Conservative 18; Mismatches 47; Indels 20; Gaps 6;
QY 25 IC-NPNCENGICLSGLADDSFSCCEPGEFAGPNCSSVVEASDEEKPPTSAGPC--IPN 81
Db 94 ICTQSPQNGGQCMYD--GGGEYHCVCILPFGHGRDC-----ERK---AGPCEQAGS 140
QY 82 PCHNGGTCISEAYRGDTFTIGYVCKCPRGFNGIHCOHININECEAPCRNGICTDLVANY 141
Db 141 PCRRGGCCODDGGFA---LNFCTCRCLVGFVGARCEVNVDDCLMRPCANGATCILDGINRF 196
QY 142 SCECPGEFMRNC 154
Db 197 SCLCPGEFAGRNC 209

RESULT 7
PCT-US01-01302-103
; Sequence 103, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01302-103

Query Match 18.3%; Score 231.5; DB 1; Length 310;
Best Local Similarity 34.6%; Pred. No. 7.7e-15;
Matches 46; Conservative 18; Mismatches 50; Indels 19; Gaps 5;

QY 24 DICNPNCENGICLSGLADDSFSCCEPGEFAGPNCSSVVEASDEEKPPTSAGPC--IPN 82
Db 101 DPCKNNCQNGGTCVQAAMLGKATCRASGFTGDC-----QYSHPCFVSXP 149
QY 83 CHNGTCISEAYRGDTFTIGYVCKCPRGFNGIHCOHININECEAPCRNGICTDLVANY 142
Db 150 CLNGTCHMLSR---DT---YECTQVGFTRKDCQWT--DACLSHPCANSTCTTANGRS 202
QY 143 SCECPGEFMRNCQ 155
Db 203 CKCLTGTGQKE 215

RESULT 8

US-60-248-505-791
; Sequence 791, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 791
; LENGTH: 4044
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-791

Query Match 16.2%; Score 204; DB 6; Length 4044;
Best Local Similarity 30.6%; Pred. No. 4e-11;
Matches 48; Conservative 11; Mismatches 52; Indels 46; Gaps 6;

QY 24 DICNPNCENGICLSGLADDS-----FSCCEPGEFAGPNCSSVY 63
Db 1249 NVCLREPCENYMKCVSLRFDSSAPPLASATLFRPIPIAGLRCPGFGTDCETEL 1308
QY 64 EVASDEEKPPTSAGPCIPNCPCHNGTCISEAYRGDTFTIGYVCKCPRGFNGIHCQ--HNIN 121
Db 1309 DLC-----NFCRRNGACARREG-----GYTCVCRPRFTGEDCELDTEAG 1347
QY 122 ECEAPECRNGICTDLV--ANYSCECP--GEFMRNCQ 155
Db 1348 RCVPGVCRNGGCTDAPNCGFRCCCPAGGAFEGPCE 1384

RESULT 9
PCT-US01-10484-63
; Sequence 63, Application PC/TUS0110484
; GENERAL INFORMATION:
; APPLICANT: Hysed, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: PCT/US01/10484
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23

```

; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 63
; LENGTH: 3014
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-10484-63

Query Match          15.5%; Score 195.5; DB 1; Length 3014;
Best Local Similarity 28.8%; Pred. No. 1.8e-10;
Matches 45; Conservative 16; Mismatches 52; Indels 43; Gaps 6;

QY 24 DICPNPCNGGICLSGLADDS-----FSCGCPGFGAPNCSSVY 63
DB 1305 NCLREPCENYMKCVSLRFDSSAPFLSSITVLFPRHPINGLRCKCPGFTGDCYETEI 1364
QY 64 EVASDEKPTSAQPCIPNCHNGTCEISEAYRGDTFTIGYCKCPRGFGNHCQNI--N 121
DB 1365 DL-----CYSDDPGANGRCRREG-----GYTCECFEDFTGEHCEVDARSG 1405
QY 122 ECEAEPCKNGICTD-LVANYSECC-PGEFMGRNCQ 155
DB 1406 RCANGVCKNGTCVNLIGGFHCVCPEGERPEYCE 1441

RESULT 10
PCT-US01-01302-90
; Sequence 90, Application PC/TUS0101302
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7207PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01302
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (195)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01302-90

Query Match          15.4%; Score 194; DB 1; Length 213;
Best Local Similarity 27.3%; Pred. No. 1.6e-11;
Matches 42; Conservative 19; Mismatches 69; Indels 24; Gaps 3;

QY 9 LTVGSLGVPOFGKGDICN-----PNPCNGGICLSGLADDSFSCPCPGFAGPNCSS 61
DB 35 LTFYLSFSLIYIKNSFCNKNTRCLNSCONNSTCKDFSKDNDSCSDTANNLDKDCDN 94
QY 62 VEVASDEKPTSAQPCIPNCHNGTCEISEAYRGDTFTIGYCKCPRGFGNHCQNIIN 121
DB 1406 RCANGVCKNGTCVNLIGGFHCVCPEGERPEYCE 1441

RESULT 11
PCT-US01-01321-1559
; Sequence 1559, Application PC/TUS0101321
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC011PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01321
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 2181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1559
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (195)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01321-1559

Query Match          15.4%; Score 194; DB 1; Length 213;
Best Local Similarity 27.3%; Pred. No. 1.6e-11;
Matches 42; Conservative 19; Mismatches 69; Indels 24; Gaps 3;

QY 9 LTVGSLGVPOFGKGDICN-----PNPCNGGICLSGLADDSFSCPCPGFAGPNCSS 61
DB 35 LTFYLSFSLIYIKNSFCNKNTRCLNSCONNSTCKDFSKDNDSCSDTANNLDKDCDN 94
QY 62 VEVASDEKPTSAQPCIPNCHNGTCEISEAYRGDTFTIGYCKCPRGFGNHCQNIIN 121
DB 95 MKD-----PCFSNPGCGSATCVNTPGER-----SFLCKCPGYSGLICETTYIG 137
QY 122 ECEAEPCKNGICTD-LVANYSECC-PGEFMGRNCQ 155
DB 138 SCGNKSCOHGICHDPIYPVICPAGYAGRFCE 171

RESULT 12
US-09-808-689-4
; Sequence 4, Application US/09808689
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Audrey Goddard
; APPLICANT: Gurney, Daniel L.
; APPLICANT: Tumas, Daniel I.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
; FILE REFERENCE: P1624R2C1
; CURRENT APPLICATION NUMBER: US/09/808,689
; CURRENT FILING DATE: 2001-03-14
```

```

DB 95 MKD-----PCFSNPGCGSATCVNTPGER-----SFLCKCPGYSGLICETTYIG 137
QY 122 ECEAEPCKNGICTD-LVANYSECC-PGEFMGRNCQ 155
DB 138 SCGNKSCOHGICHDPIYPVICPAGYAGRFCE 171

RESULT 11
PCT-US01-01321-1559
; Sequence 1559, Application PC/TUS0101321
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC011PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01321
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 2181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1559
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (195)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01321-1559

Query Match          15.4%; Score 194; DB 1; Length 213;
Best Local Similarity 27.3%; Pred. No. 1.6e-11;
Matches 42; Conservative 19; Mismatches 69; Indels 24; Gaps 3;

QY 9 LTVGSLGVPOFGKGDICN-----PNPCNGGICLSGLADDSFSCPCPGFAGPNCSS 61
DB 35 LTFYLSFSLIYIKNSFCNKNTRCLNSCONNSTCKDFSKDNDSCSDTANNLDKDCDN 94
QY 62 VEVASDEKPTSAQPCIPNCHNGTCEISEAYRGDTFTIGYCKCPRGFGNHCQNIIN 121
DB 95 MKD-----PCFSNPGCGSATCVNTPGER-----SFLCKCPGYSGLICETTYIG 137
QY 122 ECEAEPCKNGICTD-LVANYSECC-PGEFMGRNCQ 155
DB 138 SCGNKSCOHGICHDPIYPVICPAGYAGRFCE 171

RESULT 12
US-09-808-689-4
; Sequence 4, Application US/09808689
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Audrey Goddard
; APPLICANT: Gurney, Daniel L.
; APPLICANT: Tumas, Daniel I.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
; FILE REFERENCE: P1624R2C1
; CURRENT APPLICATION NUMBER: US/09/808,689
; CURRENT FILING DATE: 2001-03-14
```



```
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-689-4

Query Match
15.1%; Score 190.5; DB 5; Length 379;
Best Local Similarity 29.4%; Pred. No. 6.1e-11;
Matches 47; Conservative 15; Mismatches 47; Indels 51; Gaps 8;

QY 28 PNPENGICISGLADSFSCPEPAGPNCSSVVEASDEKPTISAGPCT-----79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 183 PGCGRNGGFC-----NERRICEPDPGFHGPHC-----EKALCTPRCMNGGLCTVPGFCICP 233
; PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHINNECEAE 126
QY 80 -----PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHINNECEAE 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 234 PGFYGVNCDKANCSTTCFNGGTC-----FYPGKCICPPGLEBGECE--ISKC-PQ 280
; PCRRNGGICITDLVANYSCCEPGEFMG-----RNCQYKCSGH 161
QY 127 PCRRNGGICITDLVANYSCCEPGEFMG-----RNCQYKCSGH 161
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 281 PCRRNGKC-----IGSKCKCKSKGYGDLCSKPYCEPGCGAH 317
; PCRRNGKC-----IGSKCKCKSKGYGDLCSKPYCEPGCGAH 317

RESULT 13
US-09-423-844-4
; Sequence 4, Application US/09423844
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Chen, Jian
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth
; APPLICANT: Pennica, Diane
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P1618P1E
; CURRENT APPLICATION NUMBER: US/09/423,844
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-423-844-4
```

```
US-09-403-296A-4
; Sequence 4, Application US/09403296A
; GENERAL INFORMATION:
; APPLICANT: Fong, Sherman
; APPLICANT: Audrey Goddard
; APPLICANT: Gurney, Austin L.
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
; FILE REFERENCE: P162AR2E
; CURRENT APPLICATION NUMBER: US/09/403,296A
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/100,858
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/19437
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-403-296A-4
```

```
Query Match
15.1%; Score 190.5; DB 5; Length 379;
Best Local Similarity 29.4%; Pred. No. 6.1e-11;
Matches 47; Conservative 15; Mismatches 47; Indels 51; Gaps 8;

QY 28 PNPENGICISGLADSFSCPEPAGPNCSSVVEASDEKPTISAGPCT-----79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 183 PGCGRNGGFC-----NERRICEPDPGFHGPHC-----EKALCTPRCMNGGLCTVPGFCICP 233
; PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHINNECEAE 126
QY 80 -----PNPCHNGTCEISEAYRGDTFIGYVCKCPRGFNGIHCOHINNECEAE 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 234 PGFYGVNCDKANCSTTCFNGGTC-----FYPGKCICPPGLEBGECE--ISKC-PQ 280
; PCRRNGGICITDLVANYSCCEPGEFMG-----RNCQYKCSGH 161
QY 127 PCRRNGGICITDLVANYSCCEPGEFMG-----RNCQYKCSGH 161
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 281 PCRRNGKC-----IGSKCKCKSKGYGDLCSKPYCEPGCGAH 317
; PCRRNGKC-----IGSKCKCKSKGYGDLCSKPYCEPGCGAH 317
```

```
RESULT 15
US-09-380-139A-4
; Sequence 4, Application US/09380139A
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC. et al.
; APPLICANT: Chen, Jian
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pennica, Diane
; APPLICANT: Wood, William I.
; APPLICANT: Yuan, Jean
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: 10466-04
; CURRENT APPLICATION NUMBER: US/09/380,139A
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US 98/19330
; PRIOR FILING DATE: 1998-09-16
; NUMBER OF SEQ ID NOS: 379
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-380-139A-4
```

```
Query Match
15.1%; Score 190.5; DB 5; Length 379;
Best Local Similarity 29.4%; Pred. No. 6.1e-11;
Matches 47; Conservative 15; Mismatches 47; Indels 51; Gaps 8;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 06:13:23 ; Search time 38.64 Seconds
(without alignments)
109,876 Million cell updates/sec

Title: US-09-237-981-29

Perfect score: 1262

Sequence: 1 MKHLVAAWLVGLSLGVPQF.....INAWTAENDRWPMIQVTG 221

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents.AA:*
2: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
7: /cgn2_6/prodata/2/1aa/backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1262	100.0	221	2	US-08-480-229C-29
2	1262	100.0	221	2	US-08-659-235C-29
3	1242	98.4	480	2	US-08-480-229C-10
4	1242	98.4	480	2	US-08-659-235C-10
5	1133.5	89.8	513	2	US-08-480-229C-14
6	1133.5	89.8	513	2	US-08-659-235C-14
7	390	30.8	463	2	US-08-162-402B-9
8	388.5	30.8	465	2	US-08-162-402B-8
9	324.5	25.7	2523	1	US-08-185-432-18
10	318	25.2	2471	1	US-08-185-432-16
11	318	25.2	2471	1	US-08-083-590A-19
12	318	25.2	2471	3	US-08-532-384-19
13	317.5	25.2	721	4	US-08-872-855-7
14	314	24.9	729	4	US-08-872-855-8
15	305.5	24.2	713	4	US-08-872-855-5
16	305.5	24.2	717	4	US-08-872-855-9
17	302.5	24.0	2703	1	US-08-185-432-19
18	298.5	23.7	321	2	US-08-480-229C-21
19	298.5	23.7	321	2	US-08-659-235C-21
20	298	23.6	1404	2	US-08-400-159-2
21	298	23.6	1404	3	US-08-611-729A-2
22	292.5	23.2	2556	1	US-08-083-590A-20
23	292.5	23.2	2556	3	US-08-532-384-20
24	291.5	23.1	720	4	US-08-872-855-4
25	291	23.1	1010	4	US-08-882-046-7
26	291	23.1	1218	4	US-08-400-159-6
27	291	23.1	1218	3	US-08-611-729A-6

28	291	23.1	1218	4	US-08-882-046-2	Sequence 2, Appl
29	290	23.0	2556	1	US-08-185-432-17	Sequence 17, Appl
30	288.5	22.9	1193	2	US-08-400-159-10	Sequence 10, Appl
31	288.5	22.9	1193	3	US-08-611-729A-10	Sequence 10, Appl
32	286	22.7	1219	4	US-08-882-046-5	Sequence 5, Appl
33	279.5	22.1	1139	1	US-08-537-210A-4	Sequence 4, Appl
34	279.5	22.1	1139	4	US-08-113-825-4	Sequence 4, Appl
35	279.5	22.1	1148	4	US-08-882-046-4	Sequence 4, Appl
36	279.5	22.1	1248	4	US-08-882-046-6	Sequence 6, Appl
37	277.5	22.0	685	4	US-08-872-855-2	Sequence 2, Appl
38	277	21.9	1064	1	US-08-537-210A-3	Sequence 3, Appl
39	277	21.9	1064	4	US-09-113-825-3	Sequence 3, Appl
40	277	21.9	1525	3	US-09-191-647-2	Sequence 2, Appl
41	275.5	21.8	387	2	US-08-162-402B-6	Sequence 6, Appl
42	271.5	21.5	833	1	US-08-264-534-6	Sequence 6, Appl
43	271.5	21.5	833	1	US-08-083-590A-2	Sequence 2, Appl
44	271.5	21.5	833	1	US-08-465-500-6	Sequence 6, Appl
45	271.5	21.5	833	2	US-08-346-126-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-480-229C-29
; Sequence 29, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quertermous, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 221 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-08-480-229C-29

Query Match 100.0%; Score 1262; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 5.2e-102; Indels 0; Gaps 0;
Matches 221; Conservative 0; Mismatches 0;

QY 181 RALFGLQKWPYAAALNKKGLINAWTAENDRMPWIOVTV 220
Db 181 RALFGLQKWPYAAALNKKGLINAWTAENDRMPWIOINL 220

RESULT 4
US-08-659-235C-10
; Sequence 10, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quertemus, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,235C
; FILING DATE: 05-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0034-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-659-235C-10

Query Match 98.4%; Score 1242; DB 2; Length 480;
Best Local Similarity 98.2%; Pred. No. 6.5e-100;
Matches 216; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKHLVAAMLLVGLSLGVQFGKGDICNPNCENGICISGLADDSFSCPEGEFAGPNC 60
Db 1 MKHLVAAMLLVGLSLGVQFGKGDICNPNCENGICISGLADDSFSCPEGEFAGPNC 60
QY 61 SVEVASDEEKPISAGPCIPNPNCHNGTCEISEAYRGDTFPGYCKCPRGFNHCOHNI 120
Db 61 SVEVASDEEKPISAGPCIPNPNCHNGTCEISEAYRGDTFPGYCKCPRGFNHCOHNI 120
QY 121 NECEAEPCRNNGICITDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
Db 121 NECEAEPCRNNGICITDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
QY 181 RALFGLQKWPYAAALNKKGLINAWTAENDRMPWIOVTV 220
Db 181 RALFGLQKWPYAAALNKKGLINAWTAENDRMPWIOINL 220
RESULT 5

US-08-480-229C-14
; Sequence 14, Application US/08480229C
; Patent No. 5874562
; GENERAL INFORMATION:
; APPLICANT: Quertemus, Thomas
; APPLICANT: Hogan, Brigid
; APPLICANT: Snodgrass, H. Ralph
; APPLICANT: Zupancic, Thomas J.
; TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
; TITLE OF INVENTION: CELL LOCUS-1
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,229C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0026-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 Pennie
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-480-229C-14

Query Match 89.8%; Score 1133.5; DB 2; Length 513;
Best Local Similarity 90.0%; Pred. No. 1.7e-90;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;
QY 1 MKHLVAAMLLVGLSLGVQFGKGDICNPNCENGICISGLADDSFSCPEGEFAGPNC 60
Db 33 MKRSVAWMLLVGLSLGVQFGKGDICNPNCENGICISGLADDSFSCPEGEFAGPNC 92
QY 61 SVEVASDEEKPISAGPCIPNPNCHNGTCEISEAYRGDTFPGYCKCPRGFNHCOHNI 120
Db 93 SVEVASDEEKPISAGPCIPNPNCHNGTCEISEAYRGDTFPGYCKCPRGFNHCOHNI 152
QY 121 NECEAEPCRNNGICITDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 180
Db 153 NECEAEPCRNNGICITDLVANYSCCEPGEFMRNCQYKCSGHLGIEGIIISNOQITASSNH 212
QY 181 RALFGLQKWPYAAALNKKGLINAWTAENDRMPWIOVTV 220
Db 213 RALFGLQKWPYAAALNKKGLINAWTAENDRMPWIOINL 253
RESULT 6
US-08-659-235C-14
; Sequence 14, Application US/08659235C
; Patent No. 5877281
; GENERAL INFORMATION:
; APPLICANT: Quertemus, Thomas
; APPLICANT: Hogan, Brigid

APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/659,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Polissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0034-999
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8664/9741
TELEX: 66141 Pennie
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235C-14

Query Match 89.88; Score 1133.5; DB 2; Length 513;
Best Local Similarity 90.08; Pred. No. 1.7e-90;
Matches 199; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 MKHLVAVMLVGLSLGVQFGKGDICNPENCGICLSGLADSFSCCEPGEFAGPNC 60
DB 33 MKRSTAVMLVGLSLGVQFGKGDICNPENCGICLSGLADSFSCCEPGEFAGPNC 92
QY 61 SVEEVAASDEEPTSGPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNI 120
DB 93 SVEEVAASDEEPTSGPCIPNCHNGTCEISEAYRGDTFYGVCCKPRGFNGIHCOHNI 152
QY 121 NCEAEPCRNCGICTDLVANYSCCEPGEFAGPNCQYKSGHGLIEGGLISNOQTASSNH 180
DB 153 NCEAEPCRNCGICTDLVANYSCCEPGEFAGPNCQYKSGHGLIEGGLISNOQTASSNH 212
QY 181 RALFGLQKWPYYALNKKGLINAWTAENDRWP-WIOVTV 220
DB 213 RALFGLQKWPYYALNKKGLINAWTAENDRWP-WIOVTV 253

RESULT 7
US-08-162-402B-9
Sequence 9, Application US/08162402B
Patent No. 5972337
GENERAL INFORMATION:
APPLICANT: CERIANI, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
APPLICANT: LARROCCA, DAVID J.
TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
TITLE OF INVENTION: GLOBULE (HMG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder & Poplawski

STREET: 444 South Flower St., 19th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Vavilana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-9

Query Match 30.98; Score 390; DB 2; Length 463;
Best Local Similarity 38.88; Pred. No. 2.5e-26;
Matches 78; Conservative 30; Mismatches 75; Indels 18; Gaps 5;

QY 23 GDICNPENCGICLSGLADSFSCCEPGEFAGPNCSSVVEASDEEPTSGPCIPNP 82
DB 25 GDCDSSLLCNGTCLTG-QDNDIYLCPEGFTGLVCE-----TERGCSPPNP 72
QY 83 CHNGTCEIS-EAYRDTFYGVCCKPRGFNGIHCOHINEAEPCRNCGICTDLVANY 141
DB 73 CYNADACIVTLDRGRDIFTEYICQCPVGSIGHCTETNYNLD--GEYFTTAVPNT 129
QY 142 SCECPGEF--MGRNCOYKSGHGLIEGGLISNOQTASSNHRLALGLQKWPYYALNKK 199
DB 130 AVPTPAFTPDLSNNLSRCSYOLGMEGALADSQISASVYMGFMGLQWGPGLARLYRT 189
QY 200 GLINAWTAENDRWPWIOVTV 220
DB 190 GIYNAWNASWYDSLFWIOVNL 210

RESULT 8
US-08-162-402B-8
Sequence 8, Application US/08162402B
Patent No. 5972337
GENERAL INFORMATION:
APPLICANT: CERIANI, ROBERTO L.
APPLICANT: PETERSON, JERRY A.
APPLICANT: LARROCCA, DAVID J.
TITLE OF INVENTION: 46 KDALTON HUMAN MILK FAT
TITLE OF INVENTION: GLOBULE (HMG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder & Poplawski
STREET: 444 South Flower St., 19th Floor
CITY: Los Angeles
STATE: CA
COUNTRY: USA

ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,402B
FILING DATE: 03-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Amzel, Viviana
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38215
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-622-7700
TELEFAX: 213-489-4210
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-162-402B-8

Query Match 30.8%; Score 388.5; DB 2; Length 465;
Best Local Similarity 40.0%; Pred. No. 3.4e-26;
Matches 88; Conservative 20; Mismatches 85; Indels 27; Gaps 6;

QY 4 LVAALLVGLSLGVPFGGDCINPNCPCNGGICLSGLADDSFSCPECPGFGPNCSSYV 63
DB 15 LCAPSLLVALD-----DFCDSSLCLNGGTCLTG-ODNDIYCLCPGCGTGLVCNE-- 62
QY 64 EVASDEKPLTSAPCIIPNCHNGGTC-EISEAYRGDTEFGYCKCPRGFGNGIHCCHNINE 122
DB 63 -----TERGICKNCHNGGCLCEELISQEVGRGVPSYCTCKGAGHCEETNY 113
QY 123 CEAPPCRNIGICTDLVANSCECPGEF--MGRNCOYKCSGHLIGEGILSNQOITASSNH 180
DB 114 YNLD--GEYMETTAVPNTAVPTPAPTPDLSNNLASKCVPELGMENGTANSQIAASSYR 170
QY 181 RALFGLOKTPYTAALNKGGLINAMTAENDRMPWIOYTV 220
DB 171 VTFGLQHWVPELARLNAGVNAWTPSSNDNPWIOYVL 210

RESULT 9
US-08-185-432-18
Sequence 18, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Bussseau, Isabelle
APPLICANT: Diederich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2523 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-185-432-18

Query Match 25.7%; Score 324.5; DB 1; Length 2523;
Best Local Similarity 32.3%; Pred. No. 7.6e-20;
Matches 64; Conservative 14; Mismatches 49; Indels 71; Gaps 6;

QY 24 DICNPNCENGICLSGLADDSFSCPECPGAGNCSSYVEVASDEKPLTSAPGCPINPC 83
DB 908 DDQCPNCPHNGSCSDGI--NMFPCNCPAGFRGKCEEDI-----NECASNPC 953
QY 84 HNGG-----CEISAYRGDPFG-----YVCKCPRGFN 112
DB 954 KNGANGCDVNSTCTCOPGSGIHCSNTPDCTESSCFNGGICIDGINFTGQCPGFT 1013
QY 113 GTHCOHINCEAPPCRNIGICTDLVANSCECPGEPMGRNQ----- 155
DB 1014 GSYCOHIDNECDSPCLNGGTCQDSYGYKCTCPQYTGILNCQVLVWCDSPCKNGKC 1073
QY 156 -----YKC---SGHLGI 164
DB 1074 WQTNFFYCECKSGMTGY 1091

RESULT 10
US-08-185-432-16
Sequence 16, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Bussseau, Isabelle
APPLICANT: Diederich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432

```

1      FILING DATE: 21-JAN-1994
2      CLASSIFICATION: 530
3      ATTORNEY/AGENT INFORMATION:
4      NAME: Mistrock, S. Leslie
5      REGISTRATION NUMBER: 18,872
6      REFERENCE/DOCKET NUMBER: 7326-006
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE: (212) 790-0900
9      TELEFAX: (212) 869-8864/9741
10     TELEEX: 66141 PENNIE
11     INFORMATION FOR SEQ ID NO: 16:
12     SEQUENCE CHARACTERISTICS:
13     LENGTH: 2471 amino acids
14     TYPE: amino acid
15     TOPOLOGY: unknown
16     MOLECULE TYPE: protein
17     OS-08-185-432-16

```

Query Match	25.2%;	Score 318;	DB 1;	Length 2471;
Best Local Similarity	41.1%;	Pred. No. 2.7e-19;		
Matches	58;	Conservative	15;	Mismatches 44;
			Indels	24;
			Gaps	4;

[illegible]

RESULT 11
 US-08-083-590A-19
 Sequence 19, Application US/08083590A
 Patent No. 5786158
 GENERAL INFORMATION:
 APPLICANT: Arcaavis-Tsakonas, S. et al.
 TITLE OF INVENTION: Therapeutic And Diagnostic Methods
 TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
 TITLE OF INVENTION: Nucleic Acids
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/083,590A
 FILING DATE: 25-JUN-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Mistrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7326-015
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 8698864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2471 amino acids

```

;      TYPE: amino acid
;;     STRANDEDNESS: single
;;     TOPOLOGY: unknown
;      MOLECULE TYPE: peptide
US-08-083-590A-19

```

Query Match	25.28;	Score 318;	DB 1;	Length 2471;
Best Local Similarity	41.18;	Pred. No. 2.7e-19;		
Matches	58;	Conservative 15;	Mismatches 44;	Indels 24;
			Gaps	4.

QY	24	DICPNPCENGGJCLISLADISDSCCPEEFAPNOCSSVEVAASDEBKETSAGPICPINC	83
		: : : : : : : : : : : : : :	
Db	913	DDCLANCONGSGSCMDV--NFSCTCLPGFTGDKCQ-----TDNNECISEPC	958
QY	84	HNGGTCEISEAYRGDPFTGYGVCKCPGPNFHHOHNINECBAPCNGGICTDVLNANSC	143
		: : : : : : : : : : : : :	
Db	959	KNGTCT-----SDYVNSTCKCKQAQPDVHCENNINECTESSCENGSTCYDGINFSFC	1011
QY	144	ECRGEFMRNCQY---KCSGH	161
		: : : : : : : : : : : : :	
Db	1012	LCPVGTGTSFCLHINECCSH	1032

```

1      RESULT 12
2      US-08-532-384-19
3      : Sequence 19, Application US/08532384
4      : Patent No. 6083904
5      : GENERAL INFORMATION:
6      :   APPLICANT: Artavanis-Tsakonas, S. et al.
7      :   TITLE OF INVENTION: Therapeutic and Diagnostic Methods
8      :   TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
9      :   TITLE OF INVENTION: Nucleic Acids
10     :   NUMBER OF SEQUENCES: 21
11     :   CORRESPONDENCE ADDRESS:
12     :     ADDRESSEE: Pennie & Edmonds
13     :     STREET: 1155 Avenue of the Americas
14     :     CITY: New York
15     :     STATE: New York
16     :     COUNTRY: U.S.A.
17     :     ZIP: 10036
18     :   COMPUTER READABLE FORM:
19     :     MEDIUM TYPE: Floppy disk
20     :     COMPUTER: IBM PC compatible
21     :     OPERATING SYSTEM: PC-DOS/MS-DOS
22     :     SOFTWARE: PatentIn Release #1.0, Version #1.25
23     :   CURRENT APPLICATION DATA:
24     :     APPLICATION NUMBER: US/08/532,384
25     :     FILING DATE:
26     :       CLASSIFICATION: 424
27     :       PRIOR APPLICATION DATA:
28     :         APPLICATION NUMBER: 08/083,590
29     :         FILING DATE: 25-JUN-1993
30     :       ATTORNEY/AGENT INFORMATION:
31     :         NAME: Mastro, S. Leslie
32     :         REGISTRATION NUMBER: 18,872
33     :         REFERENCE/DOCKET NUMBER: 7346-015
34     :       TELECOMMUNICATION INFORMATION:
35     :         TELEPHONE: 212 790-9090
36     :         TELEFAX: 212 8698864/9741
37     :         TELEX: 66141 PENNIE
38     :       INFORMATION FOR SEQ ID NO: 19:
39     :         SEQUENCE CHARACTERISTICS:
40     :           LENGTH: 2471 amino acids
41     :           TYPE: amino acid
42     :           STRANDEDNESS: single
43     :           TOPOLOGY: unknown
44     :         MOLECULE TYPE: peptide
45     :         US-08-532-384-19

```

Query Match	25.2%;	Score 318;	DB 3;	Length 2471;
Best Local Similarity	41.1%;	Pred. No. 2.7e-19;		

